

A new Bismut-Elworthy-Li-formula
for diffusions with singular coefficients
driven by a pure jump Lévy process
and applications to life insurance

by

Tor Martin Christensen

Thesis
for the degree of
Master of science

(Master i Modelling og Dataanalyse)



Faculty of Mathematics and Natural Sciences
University of Oslo

April 2015

Det matematisk- naturvitenskapelige fakultet
Universitetet i Oslo

Abstract

The main result of my mine in the master thesis is a new Bismut-Elworthy-Li formula with respect to a pure jump Lévy noise driven stochastic differential equation (SDE), with non-Lipschitz continuous coefficients. More precisely, I obtain in this thesis for the first time the following representation:

$$\frac{\partial}{\partial x} E[g(X_T^x)] = E \left[g(X_T^x) \cdot \frac{1}{S_T^{\alpha/2}} \int_0^T \frac{\partial}{\partial x} X_s^x dL_s \right],$$

where g is a continuous function and where X_t^x satisfies the SDE

$$X_t^x = x + \int_0^t b(X_s^x) ds + L_t, \quad 0 \leq t \leq T$$

for an α -stable process L_t , $\alpha \in (1, 2)$ and a $\frac{\alpha}{2}$ -stable subordinator $S_t^{\alpha/2}$. Here we only require that the drift coefficient is Hölder continuous. We mention that the above result, which is a generalization of the paper [17] to the case of singular drift coefficients b , was first obtained in [12] for SDE's driven by Brownian motion.

Let me also remark that the above formula can be considered a representation of "pure jump Lévy" delta of a financial claim $h = g(X_T^x)$ with an underlying asset price dynamics given by X_s , $0 \leq s \leq T$, which does not involve a derivative of the payoff function g .

This thesis consists of 5 chapters, where chapter 1 is an introduction to what Greeks are and why they are interesting in finance. In chapter 2 there is an overview and discussion of basic methods for the calculation of Greeks in the literature. In chapter 3 there is an implementation of what we refer to as Zhang's formula, namely a Bismut-Elworthy-Li type formula. This is a "derivative free" type formula for SDEs driven by pure jump process, namely an α -stable process. In the first part of chapter 3 simulations are conducted confirming that Zhang formula in numerical implementations works, then there is presented an application of this formula to life insurance, where we also conduct simulations.

Chapter 4 is the highlight of this thesis, where we derive a Bismut-Elworthy-Li type formula for the Greek Delta. This derivative free representation is obtained by using methods in [17] and [8]. The formula can be regarded as an extension of Zhang's formula in case of the Greek Delta, in the sense that we deal with Hölder coefficients and don't demand that the coefficients have continuous first order derivative.

Chapter 5 suggests possible extensions to this thesis.

Acknowledgements

I would like to thank my supervisor Frank Proske, for providing me with a challenging and very interesting topic for my thesis, along with his help and guidance.

I want to thank my parents Tor and Mette for their support during my studies. I would like to thank my fellow students at room 802 for an enjoyable working environment these past two years. I want to thank Maja, and especially Lars for interesting discussions and finding misprints. I would like to thank family and friends for encouragement, especially Jonas Skaalen.

And last but not least my girlfriend, Marte, for all her love and invaluable support.

Notation

The space $L^2(P)$ is the space of square integrable random variables, the norm of a random variable $X \in L^2(P)$ is given by

$$\|X\|_{L^2(P)} := (E[X^2])^{\frac{1}{2}} = \left(\int_{\Omega} X^2(\omega) dP(\omega) \right)^{\frac{1}{2}}. \quad (1)$$

In the general case, random variables $X \in L^p(P), p \in [1, \infty)$ are equipped with the norm

$$\|X\|_{L^p(P)} = \left(\int_{\Omega} |X(\omega)|^p dP(\omega) \right)^{\frac{1}{p}}$$

The indicator function is defined as

$$\mathbf{1}_A = \mathbf{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

When an equation is referred to, it goes for each chapter, e.g. equation (2.3) refers to equation 3 in chapter 2, also theorem's, lemma's etc. are numbered per chapter. References are noted by [], e.g. [2] is references number 2 listed in the bibliography.

Contents

Preface	i
Abstract	i
Acknowledgements	iii
Notation	v
1 Introduction	1
2 Overview of basic methods for calculation of Greeks	5
2.1 Overview of numerical techniques for computations of Greeks .	6
2.1.1 The finite difference method	9
2.1.2 Numerical method through Malliavin Calculus	10
2.1.3 The Likelihood Ratio Method	12
2.2 Overview of concepts in Malliavin calculus in case of the Brownian motion	13
2.3 Malliavin calculus in case of Lévy processes	19
3 Implementation of Greeks driven by α-stable processes	23
3.1 Lévy processes	25
3.2 Modeling with the α -stable processes	29
3.3 Derivative formula and gradients estimates for SDE's	33
3.4 Simulation of the derivative of a caplet with respect to the initial interest rate, under the Vasicek interest rate model . . .	35
3.4.1 The Vasicek interest rate model	35
3.4.2 Simulation of the derivative of a caplet with respect to the initial interest rate	38
3.5 Application to life insurance	42
3.5.1 Framework in life insurance	43
3.5.2 Unit-linked policies	46
3.5.3 Simulation of the sensitivity of unit-linked policies . . .	48

4	Derivation of the delta in the case of SDE's with Hölder drift coefficients	55
4.1	Framework	56
4.2	Properties of an SDE with Hölder coefficients	60
4.3	A Bismut-Elworthy-Li formula ("delta") with Hölder coefficients	66
5	Extensions	73
A	Appendix	75
A.1	Calculations	75
A.2	R-code for simulations	78
A.3	Bibliography	104

Chapter 1

Introduction

In the world of finance, there are numerous types of contracts often known as financial derivatives. The price of such contracts are derived from the underlying asset, e.g. stocks, bonds, interest rates, currencies. A well known type of contract is the European call option. More precisely let $S(T)$ denote the value of the underlying asset, where T is the time to maturity, that is, when the option can be exercised. Furthermore, if K denotes the strike price, the option takes the form $\max(S(T) - K, 0)$, where the investor pays an agreed upon sum to the other party when the contract starts. This gives the investor the right to purchase the underlying asset at a price in the future agreed upon today. Such contracts can be used as an insurance, in the sense that an investor can buy protection if the value of the underlying asset the investor holds crosses a threshold. This strategy is a type of hedge, that is to reduce the risk. A highly interesting topic is how sensitive they are when a parameter changes, maybe the underlying asset becomes more volatile or the drift changes. What if the value of the underlying asset changes? This branch of financial mathematics is known as sensitivity analysis or more commonly referred to as Greeks. This tool is often applied by investors in the financial market, as risk measure, used to hedge their positions.

When one computes Greeks in finance, one investigates the market sensitivities of financial derivatives (e.g. call option, put option, digital option etc.) when parameters in a given model change. These quantities are often denoted by Greek letters, hence the name Greeks. To obtain a Greek the main idea is to take the derivative of the risk-neutral price of an option (e.g. call option) with respect to the parameter one is interested in. More precisely, if we let

$$V = E_Q[e^{-\int_0^T r(s)ds} \phi(S(T))]$$

denote the risk-neutral price, where ϕ denotes the payout function, $S(T)$ the

value of the underlying asset (e.g. a stock) at terminal time T . Furthermore r denotes the overnight interest rate (so $e^{-\int_0^T r(s)ds}$ is the discount factor), and the expectation is taken with respect to the risk neutral probability measure Q . To obtain a Greek, one must take the partial derivative of a parameter of the risk-neutral price, e.g.:

- Delta is used to construct the delta hedge in a portfolio, denoted $\Delta = \frac{\partial V}{\partial x}$. Delta measures the sensitivity of change in the price x of the underlying asset. In fact taking the derivative with respect to the underlying asset gives us the hedge ratio, which is needed to obtain the replicating portfolio.
- Gamma is the derivative of the delta with respect to the price x of the underlying asset, $\Gamma = \frac{\partial^2 V}{\partial x^2}$.
- Rho measures the sensitivity to the interest rate, which is obtained by taking the derivative with respect to r , $\rho = \frac{\partial V}{\partial r}$.
- Theta is obtained by taking the derivative with respect to time: $\theta = -\frac{\partial V}{\partial T}$, theta measures the sensitivity to the time to maturity.
- Vega, which is not a greek letter (but denoted by the Greek letter ν), measures sensitivity of V , with respect to the volatility of the underlying asset: $\nu = \frac{\partial V}{\partial \sigma}$.

These are some of the most common Greeks, but the possibilities are endless. Where the latter statement entails that one can find high order of Greeks, that is, to take higher order derivatives of a risk-neutral price (V). We will consider first order Greeks in this thesis. Delta is a very interesting Greek, it is used to obtain the hedge ratio, which is needed to find the replicating portfolio of a financial derivative, such as the call option. Where the replicating portfolio of an option is the portfolio strategy needed to produce the same outcome as the option. In a replicating portfolio of an option, one invest in the underlying asset of the option and the bank. Taking the derivative of a risk-neutral price to e.g. obtain one of the Greeks above, can be accomplished in a straight forward manner under nice conditions, that is when one is allowed to commute differentiation and expectation. Under the assumptions of a Black-Scholes model Greeks are relatively straight forward to compute.

However, in general it is often impossible to obtain an analytical expression for a Greek. Hence one would resort to numerical methods to obtain the Greek that one sets out to find. For instance, if the payout function

is discontinuous it may be impossible to obtain the derivative. One could resort to the so-called density method, where one moves the derivative of the payout function to only depend on the density function¹. This method works well, but one is required to have an explicit expression of the density function. In fact it turns out that this method is a special case of the general Malliavin approach. With tools from Malliavin calculus, it is possible to obtain a derivative free form of Greeks by

$$E_Q[e^{-\int_0^T r(s)ds} \phi(S(T))\pi] \quad (1.1)$$

where π is the so-called Malliavin weight. This was done by the authors in [6].

First we will look into the Malliavin weight π . To make this technique less mysterious we include an overview of some important concepts of Malliavin calculus in the continuous case, namely Brownian motion. Readers familiar with basic methods for numerical approximation of Greeks, such as finite difference, likelihood ratio method and the application of Malliavin calculus, may skip chapter 2.

After an overview of basic methods to obtain Greeks we will introduce the concept of Lévy processes. We look at Lévy processes or more precisely pure jump processes, as a model able to capture jumps is considered more realistic, e.g. if there is an abrupt change in a country's monetary policy. In fact, we will study pure jump processes, that is, we assume that the process only consists of jumps. More precisely we will work with the α -stable process, which is explained in chapter 3. With help from a Bismut-Elworthy-Li type formula developed in [17] we will simulate/approximate Greeks, ran by α -stable processes.

¹See the section about the Likelihood Ratio Method for details on this technique.

Chapter 2

Overview of basic methods for calculation of Greeks

Obtaining Greeks analytically under a framework of differentiable payoff functions and continuous process can be a straight forward process. When we have nice conditions that allows us to commute differentiation and expectation. On the other hand, it can be quite challenging in certain scenarios, where one has to deal with jump processes, non-differentiable payoff functions or very complicated options. One can approximate solutions numerically by various techniques, using Monte Carlo simulation, which we will see can be used to obtain quite accurate estimates. In this chapter we will first get an overview of techniques used to obtain Greeks numerically, where the application from Malliavin calculus is included. Furthermore because of the Malliavin weight π there is included an overview of some important concepts in Malliavin calculus.

We will frequently encounter the stochastic process Brownian motion throughout this thesis, which has the properties:

Definition 2.1. *The Brownian motion $B(t)$ is a continuous ¹ stochastic process which have the following properties*

- i) Independent increments: The random variable $B(t) - B(s)$ is independent of the random variable $B(u) - B(v)$ where $t > s \geq u > v \geq 0$.*
- ii) Normal increments: The distribution of $B(t) - B(s)$ for $t > s \geq 0$ is normal with expectation 0 and variance $t - s$.*

¹But nowhere differentiable.

Also, we have the definition of what a σ -algebra is, which we will encounter throughout the thesis. We will use \mathcal{F}_t as the smallest σ -algebra generated by the Brownian motion up to time t .

Definition 2.2. *If Ω is a given set, then a σ -algebra \mathcal{F} on Ω is a family \mathcal{F} of subsets of Ω with the following properties:*

- (i) $\emptyset \in \mathcal{F}$
- (ii) $F \in \mathcal{F} \Rightarrow F^C \in \mathcal{F}$, where $F^C = \Omega \setminus F$ is the complement of F in Ω
- (iii) $A_1, A_2, \dots \in \mathcal{F} \Rightarrow A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

2.1 Overview of numerical techniques for computations of Greeks

If we assume that we have an underlying asset described by the stochastic differential equation (SDE)

$$dX(t) = b(X(t))dt + \sigma(X(t))dB(t), \quad (2.1)$$

where the coefficients b and σ are Lipschitz continuous, i.e. satisfies the usual conditions to make sure that the solution of eq. (2.1) exist and is unique ². Furthermore $\{B(t), 0 \leq t \leq T\}$ is the Brownian motion with values in \mathbb{R}^n . Then the solution $\{X(t); 0 \leq t \leq T\}$ is a Markov process with values in \mathbb{R}^n . Let

$$u(x) = E[\phi(X(T)) | X(0) = x], \quad (2.2)$$

where T is the maturity time of an option and ϕ is the payoff function, e.g. call option, digital option, Asian option. The function $u(x)$ denotes the price of the option (with interest rate $r = 0$), which can be computed by Monte Carlo methods. One can investigate how sensitive an option is with respect to its different parameters. One needs to compute the differentials of $u(x)$ with respect to the parameters one is interested in, e.g. the drift coefficient b , the volatility σ or the initial value x . Finding Greeks analytically may sometimes be impossible, thus one needs to resort to numerical methods. To be able to understand the numerical methods, in the next subsections, we will first see how to simulate an expectation and an arbitrary dynamics of the form (2.1)

²See theorem 5.2.1 in [4] for details on the conditions concerning existence and uniqueness of SDE's.

The approximation in the Monte Carlo simulation is found by computing the payoff function m times, and weighting each simulation equally. The estimate for the expectation becomes more accurate as the number of simulations increases. As an example of a process of type (2.1), take the geometric Brownian motion, which is described by the dynamic

$$dX(t) = \mu X(t)dt + \sigma X(t)dB_t \quad (2.3)$$

for given constant drift and volatility respectively denoted by $\mu, \sigma \in \mathbb{R}$. The solution of the dynamic takes the form (proof is given in Lemma A.2)

$$X_t = x \cdot \exp \left(\left(\mu - \frac{1}{2}\sigma^2 \right)t + \sigma B_t \right), \quad x = X(0), \quad (2.4)$$

which can be obtained by using the Itô formula on eq. (2.3). If one were to apply this in practice, e.g. take the derivative with respect to the underlying asset (obtain the hedge ratio), then one need to apply Girsanov's theorem, to make sure that the process is risk neutral, so there are no arbitrage opportunities³.

If we want to approximate eq.(2.2) numerically, where we let X_T be the geometric Brownian motion, i.e. eq.(2.4).

The process varies in the sense that we need to simulate a new B_t for each run, which will yield a different path for each simulation. This is emphasized by the subscript j in eq. (2.7). First one need the simulations of X_t , which can be found recursively; for $n \in \mathbb{N}$ let $0 = t_0 \leq \dots \leq t_n = T$ be uniformly distributed, so that the n time points have the same distance between them, denoted by Δt which becomes $\Delta t = \frac{T}{n}$. Then we can simulate $B_t = (B_{t_0}, B_{t_1}, \dots, B_{t_n})$ by drawing $N(0, \Delta t)$ ($N(\mu, \sigma^2)$ denotes the normal distribution with expectation μ and variance σ^2) as follows:

$$\begin{aligned} B_{t_0} &= Y_0 \\ B_{t_1} &= B_{t_0} + Y_1 \\ B_{t_2} &= B_{t_1} + Y_2 \\ &\vdots \\ B_{t_n} &= B_{t_{n-1}} + Y_n \end{aligned}$$

Where Y_0, Y_1, \dots, Y_n are i.i.d $N(0, \Delta t)$. Utilizing this and starting the recur-

³For more about risk neutrality see [1].

sion at the point $X_{t_0} = x$ we get the following:

$$X_{t_n} = X_{t_{n-1}} \cdot \exp \left(\left(\mu - \frac{1}{2} \sigma^2 \right) \Delta t + \sigma (B_{t_n} - B_{t_{n-1}}) \right) \quad (2.5)$$

$$= x \cdot \exp \left(\left(\mu - \frac{1}{2} \sigma^2 \right) T + \sigma B_{t_n} \right), \quad (2.6)$$

where μ and σ are given constants, and x is the initial value of the underlying asset. We see that the geometric Brownian motion has the Markov property, which simplifies computation; thus, conducting a numerical simulation on a computer, eq. (2.6) is preferable to eq. (2.5) as this does not require any looping. Furthermore, one needs to simulate an expectation, which we will do by weighting each simulation equally, so we obtain the average, hence we need to simulate equation (2.6) m times. We let the simulations (end point of each recursion) be put into one vector, more precisely we let

$$\mathbf{X} = (X_1(T), X_2(T), \dots, X_m(T))$$

denote the m simulations, where ⁴ \mathbf{X}_j is an arbitrary simulation ($1 \leq j \leq m$) We can then use the following approximation:

$$u(x) = E[\phi(X(T))] \approx \frac{1}{m} \sum_{j=1}^m \phi(\mathbf{X}_j(T)). \quad (2.7)$$

where m is the number of simulations, higher values of m yields a more accurate value for the expectation. Since equation (2.7) is an approximation one need to consider the trade-off between increasing the accuracy for the expectation and the computation time, when conducting simulations.

If one encounter a complex dynamic, which can be hard or even impossible to solve analytically, one can solve the problem numerically. A straight forward approximation is the Euler scheme. More precisely, if one has a general dynamic of the form

$$dX(t) = b(X(t))dt + \sigma(X(t))dB(t)$$

(as earlier), then one can utilize the following approximation to find a realization of $X(T)$ numerically:

⁴The parameters of the process \mathbf{X}_j are suppressed, so when we use an illustrating example later on, where we take the derivative with respect to a arbitrary parameter, it's assumed that the process is a function of that parameter.

- Let $0 = t_0 \leq \dots \leq t_n = T$ be uniformly distributed, so that the n time points have the same distance between them. Thus $\Delta t = \frac{T}{n}$.
- For $0 \leq i \leq n$ simulate the recursion

$$X_{i+1} = b(X_i)\Delta t + \sigma(X_i)\Delta B_i,$$

where $X_0 = x$ and

$$\Delta B_i = B_{t_{i+1}} - B_{t_i}.$$

2.1.1 The finite difference method

A basic method for computing the sensitivity of an option, e.g. delta, gamma, rho, is to deploy the finite difference method, which is based on finding an approximation of the derivative numerically. If we for instance look at the Delta, one first have to use Monte Carlo to obtain an estimate of eq. (2.2) and an estimate for $u(x + \epsilon)$ for a small $\epsilon > 0$. Using the forward finite difference estimator, one gets the estimate

$$\Delta = \frac{\partial u(x)}{\partial x} \approx \frac{u(x + \epsilon) - u(x)}{\epsilon}. \quad (2.8)$$

Alternatively, one can use the central difference method to obtain an even better approximation:

$$\Delta = \frac{\partial u(x)}{\partial x} \approx \frac{u(x + \epsilon) - u(x - \epsilon)}{2\epsilon}. \quad (2.9)$$

With this method we get a faster rate of convergence.

In fact we can use eq.(2.9) once more to obtain an approximation for the Gamma:

$$\Gamma = \frac{\partial^2 u(x)}{\partial x^2} \approx \frac{u(x + \epsilon^*) - 2u(x) + u(x - \epsilon^*)}{\epsilon^{*2}} \quad (2.10)$$

where ⁵ $\epsilon^* = 2\epsilon$. One could argue here that since we pick an arbitrary ϵ , that 2ϵ can easily be replaced by a new ϵ since all we demand is that $\epsilon > 0$. The reason why we use ϵ^* is because of the choice of ϵ in eq. (2.9) and eq. (2.10) isn't necessarily the same. The choice of ϵ can't be too big or too small ⁶. Preferably one should use an equation which can provide us with the optimal choice of ϵ . For the choice of ϵ in the case of Gamma see eq.(11.8) in [10].

⁵See lemma A.4 on the derivation of the Gamma approximation.

⁶For a details on this topic see e.g. [10] page 141.

Approximation with Monte Carlo makes the central difference method, in the case of the delta take the form

$$\Delta \approx \frac{1}{2m\epsilon} \sum_{j=1}^m [\phi(\mathbf{X}_j(x + \epsilon)) - \phi(\mathbf{X}_j(x - \epsilon))]. \quad (2.11)$$

The numerical approximation for Gamma takes the form

$$\Gamma \approx \frac{1}{m\epsilon^2} \sum_{j=1}^m [\phi(\mathbf{X}_j(x + \epsilon^*)) - 2 \cdot \phi(\mathbf{X}_j(x)) + \phi(\mathbf{X}_j(x - \epsilon^*))] \quad (2.12)$$

It's been proved (see [6]) that the convergence rate of the forward finite difference estimator ⁷ is $n^{-1/4}$. By using the central difference estimator, one has a convergence rate of $n^{-1/3}$, which significantly decrease the computation time. It's even possible to obtain a convergence rate of $n^{-1/2}$ if one use common random numbers (a variance reduction technique) for both estimators. A drawback of the finite difference method is that it lacks the ability to deal with non-differentiable payoff functions, such as the digital option which pays zero or one: $\phi(X(T)) = \mathbb{1}_{\{X(T) > K\}}$ (where K denotes the strike price of the option). By means of Malliavin calculus one can overcome this obstacle, namely be able to approximate a Greek that has a discontinuous payoff function. In fact the method in the next subsection can also be deployed if one has to deal with a Lévy process, that is, when we allow for jumps in a stochastic process.

2.1.2 Numerical method through Malliavin Calculus

An important application from Malliavin Calculus is that one can obtain closed theoretical formulas for Greeks, where the differentiation operator is "moved away" from the payoff function. This method has the ability to deal with jump and continuous processes, where we don't need to know the explicit density ⁸. These formulas might be theoretically difficult to solve, but by deploying Monte Carlo simulation it's possible to obtain accurate approximations of Greeks. More precisely, a numerical approximation using Malliavin calculus in combination with Monte Carlo yields a convergence

⁷Under the assumption that both estimators are drawn independently of each other in the Monte Carlo simulation.

⁸If one do however have the explicit density the Malliavin weight is easy to find, as this mean we are essentially using the likelihood ratio method, as we shall see in the forthcoming subsection.

rate of $n^{-1/2}$. As shown by Fournié et al. in [6] by using Malliavin calculus one can express Greeks in the following way:

$$E[\pi\phi(X(T))|X(0) = x],$$

where π is the Malliavin weight. An important feature here is that π does not depend on the payoff function ϕ . The authors in [6] conducted numerical experiments showing how well this works and compared it with the finite difference method. This was done in a setting where they already knew the theoretical values of a given Greek. Their numerical experiments showed that this application of Malliavin calculus is very useful to compute Greeks. Note that the weights used were not unique. In the first paper they chose arbitrary non-complicated weights, where they in a follow up paper, namely [7] discuss the choice of the optimal Malliavin weight. They investigate the optimal choice of weight in the sense of minimal variance. More precisely, for all possible weights π the idea is to minimize

$$\begin{aligned} V(\pi) &= E[|\phi(X(T))\pi - \frac{d}{d\theta}u(x)|^2] \\ &= E[\phi^2(X(T))\pi^2] - E[\phi(X(T))\pi_0]^2 \\ &= E[\phi^2(X(T))(\pi^2 - \pi_0^2)] + V(\pi_0) \end{aligned}$$

where π_0 is the weight with the smallest variance and θ is an arbitrary parameter that the risk-neutral price $u(x)$ depend on.

The method of using the so-called Malliavin weight works well in the case of discontinuous payoff functions, but for some Greeks the computation might be slow e.g. if one has powers of the Brownian motion.⁹

Usually the finite difference method performs just as good as deploying Malliavin calculus, except when we for instance deal with non-smooth payoff functions. Thus in some cases, it's more favorable to deploy the finite difference method (depending on the payoff function) because it's more cumbersome to use Malliavin calculus in this case, as it requires that we set of heavy theoretical machinery. Note that a downside of estimation by means of Monte Carlo, is that it might converge slowly, and in some cases the estimate might be poor.

⁹The authors of [6] point out that some weights may have powers of the Brownian motion, which slows down the Monte Carlo simulations, so they introduce the idea of what they call a localized version, to improve numerical simulations. For more on this consult [6].

2.1.3 The Likelihood Ratio Method

The likelihood ratio method is a special case of an application of Malliavin calculus. Given that we are able to obtain the explicit density function (which must depend on the parameter we are differentiating with respect to) for the payoff function $\phi(X(T))$, one can move the differentiation operator inside the expectation. Thus we can move the dependence from the payoff function to the density function, which makes us able to treat non-smooth payoff functions. Letting the risk-neutral price of the option be on the form $u(x) = E[\phi(X(T))|X(0) = x]$, where $\phi(X(T))$ denotes the payoff function. We can observe how one can move the differentiation to only depend on the density function of the payoff function:

$$\begin{aligned}\frac{\partial}{\partial z} E[\phi(X(T))] &= \frac{\partial}{\partial z} \int \phi(y) f_z(y) dy \\ &= \int \phi(y) \frac{\partial}{\partial z} f_z(y) dy \\ &= \int \phi(y) \frac{\partial}{\partial z} f_z(y) \cdot \frac{f_z(y)}{f_z(y)} dy \\ &= \int \left(\phi(y) \frac{\partial}{\partial z} \log(f_z(y)) \right) f_z(y) dy \\ &= E[\phi(X(T)) \frac{\partial}{\partial z} \log(f_z(y))]\end{aligned}$$

where z is some arbitrary parameter that our payoff function depend on, and $f_z(y)$ denotes the density function of the payoff function. The name of this method originates from the fact that the term

$$\frac{\frac{\partial}{\partial z} f_z(y)}{f_z(y)} = \frac{\partial}{\partial z} \log(f_z(y))$$

in the above equality, could be regarded as a likelihood ratio in the sense that its the ratio between two density functions. Computing the Delta (done in [10]) by means of Monte Carlo, in the case of the likelihood ratio method one obtains the approximation

$$\Delta \approx \frac{1}{m} \sum_{j=1}^m [\phi(\mathbf{X}_j(T)) \cdot \frac{\partial}{\partial z} \log(f_z(\phi(\mathbf{X}_j(T))))]$$

where we simulate m times in order to compute the expectation.

2.2 Overview of concepts in Malliavin calculus in case of the Brownian motion

In the following section we will look at basic concepts of Malliavin calculus as it is presented by the authors in [4]. This will make us more familiar with the Malliavin weight π . Malliavin calculus was introduced by Paul Malliavin, and was used as a tool to study the smoothness of densities of solutions of stochastic differential equations driven by Brownian motion. At first this calculus was considered complicated with limited applications, but it soon became clear that Malliavin calculus was significant, due to applications discovered in stochastic control, insider trading and the application in sensitivity analysis. The application in sensitivity analysis was discovered by Fournié et al. ([6]), where they were able to move the derivative from the payoff function, which means one can treat a great deal of complicated options (which could be non-differentiable). The so-called Malliavin weight π is what we shall embark on in the following section. Here we will get an overview of some of the important concepts used in the derivation of the weight π .

The Wiener-Itô chaos expansion

The weight π is defined through the Malliavin derivative, which is defined through the Wiener-Itô chaos expansion. We first need the notion of the n -fold iterated Itô integral, in which the Wiener-Itô chaos expansion is defined through. Let $B(t) = B(\omega, t)$ ($B(0) = 0$), $\omega \in \Omega$, $t \in [0, T]$ ($T > 0$) be the Brownian motion on the complete probability space (Ω, \mathcal{F}, P) . Also we denote by \mathcal{F}_t the σ -algebra generated by $B(s)$, $0 \leq s \leq t$. We have the following definition of a symmetric function:

Definition 2.3. *A real function $g : [0, T]^n \rightarrow \mathbb{R}$ is called symmetric if*

$$g(t_{\sigma_1, \dots, t_{\sigma_n}}) = g(t_1, \dots, t_n)$$

for all permutations $\sigma = (\sigma_1, \dots, \sigma_n)$ of $(1, 2, \dots, n)$

The symmetrization \tilde{f} of a real function f on $[0, T]^n$ is defined by

$$\tilde{f}(t_1, \dots, t_n) = \frac{1}{n!} \sum_{\sigma} f(t_{\sigma_1}, \dots, t_{\sigma_n})$$

where we take the sum over all permutations of σ . Furthermore, we will work in the space of square integrable Borel real functions on $[0, T]^n$ denoted by

$L^2([0, T]^n)$, where the norm is defined as

$$\|g\|_{L^2([0, T]^n)}^2 = \int_{[0, T]^n} g^2(t_1, \dots, t_n) dt_1 \cdots dt_n < \infty \quad (2.13)$$

Let $\tilde{L}^2([0, T]^n)$ denote the space of symmetric square integrable functions on $[0, T]^n$, which is a subspace of $L^2([0, T]^n)$.

Let $S_n = \{(t_1, \dots, t_n) \in [0, T]^n : 0 \leq t_1 \leq t_2 \leq \dots \leq t_n \leq T\}$. Then we have the following definition of the n -fold iterated Itô integral:

Definition 2.4. *Let f be a deterministic function defined on S_n ($n \geq 1$) such that*

$$\|f\|_{L^2(S_n)}^2 := \int_{S_n} f^2(t_1, \dots, t_n) dt_1 \cdots dt_n < \infty$$

Then we can define the n -fold iterated Itô integral as

$$J_n(f) := \int_0^T \int_0^{t_n} \cdots \int_0^{t_3} \int_0^{t_2} f(t_1, \dots, t_n) dB(t_1) dB(t_2) \cdots dB(t_{n-1}) dB(t_n).$$

The Wiener-Itô chaos expansion is a way of representing square integrable random variables, namely variables $X \in L^2(P)$, which is defined through symmetric functions. We have the following definition, of the so-called n -fold iterated Itô integrals

Definition 2.5. *If $g \in \tilde{L}^2([0, T]^n)$ we define*

$$I_n(g) := \int_{[0, T]^n} g(t_1, \dots, t_n) dB(t_1) \cdots dB(t_n) := n! J_n(g)$$

we also call n -fold iterated Itô integrals the $I_n(g)$ here above.

In practice one can use Hermite polynomials to obtain the iterated Itô integrals, which relies on the relationship between the Hermite polynomials and the Gaussian distribution density. The Hermite polynomials $h_n(x)$ are defined as

$$h_n(x) = (-1)^n e^{\frac{1}{2}x^2} \frac{d^n}{dx^n} (e^{-\frac{1}{2}x^2}), \quad n = 0, 1, 2, \dots,$$

In fact one can obtain the iterated Itô integral by the following formula:

$$n! \int_0^T \int_0^{t_n} \cdots \int_0^{t_2} g(t_1) g(t_2) \cdots g(t_n) dB(t_1) \cdots dB(t_n) = \|g\|^n h_n \left(\frac{\theta}{\|g\|} \right),$$

where $\|g\| = \|g\|_{L^2([0, T])}$ and $\theta = \int_0^T g(t) dB(t)$.

Example 2.6. Let $g \equiv 1$ and $n = 3$, we then have that

$$6 \int_0^T \int_0^{t_3} \int_0^{t_2} 1 dB(t_1) dB(t_2) dB(t_3) = T^{3/2} h_3 \left(\frac{B(T)}{T^{1/2}} \right) = B^3(T) - 3TB(T).$$

The Wiener-Itô chaos expansion is a way of representing random variables in $L^2(P)$ through the n -fold iterated Itô integrals, more precisely:

Theorem 2.7. *Let ξ be an \mathcal{F}_t -measurable random variable in $L^2(P)$. Then there exists a unique sequence $\{f_n\}_{n=0}^\infty$ of functions $f_n \in \tilde{L}^2([0, T]^n)$ such that*

$$\xi = \sum_{n=0}^{\infty} I_n(f_n),$$

where the convergence is in $L^2(P)$.

Proof. See proof of Theorem 1.10 in [4] □

There are different ways of defining the Malliavin derivative. In this brief summary we will see how it is defined through the Wiener-Itô chaos expansion. Let's take a small detour mentioning its adjoint operator, namely the Skorohod integral.

The Skorohod integral

The Skorohod integral is an extension of the Itô integral, in the sense that under certain assumptions the two integrals will coincide. The Skorohod integral is defined through the Wiener-Itô chaos expansion:

In the following definition we assume that the chaos expansion of a random variable $u(t)$ is of the form

$$u(t) = \sum_{n=0}^{\infty} I_n(f_{n,t})$$

where $f_{n,t} = f_{n,t}(t_1, \dots, t_n) = f_n(t_1, \dots, t_n, t)$ are symmetric functions in $\tilde{L}^2([0, T]^n)$.

Definition 2.8. *Let $u(t)$, $t \in [0, T]$, be a measurable stochastic process such that for all $t \in [0, T]$ the random variable $u(t)$ is \mathcal{F}_T -measurable and $E[\int_0^T u^2(t) dt] < \infty$. Let its Wiener-Itô chaos expansion be*

$$u(t) = \sum_{n=0}^{\infty} I_n(f_{n,t}) = \sum_{n=0}^{\infty} I_n(f_n(\cdot, t)).$$

Then we define the Skorohod integral of u by

$$\delta(u) := \int_0^T u(t) \delta B(t) := \sum_{n=0}^{\infty} I_{n+1}(\tilde{f}_n) \quad (2.14)$$

when convergent in $L^2(P)$. Here \tilde{f}_n , $n = 1, 2, \dots$, are the symmetric functions derived from $f_n(\cdot, t)$, $n = 1, 2, \dots$. We say that u is Skorohod integrable and we write $u \in \text{Dom}(\delta)$ if the series in (2.14) converges in $L^2(P)$.

Here the symmetrization f_n is given by

$$\tilde{f}_n(t_1, \dots, t_{n+1}) = \frac{1}{n+1} [f_n(t_1, \dots, t_{n+1}) + f_n(t_2, \dots, t_{n+1}, t_1) + \dots + f_n(t_1, \dots, t_{n-1}, t_{n+1}, t_n)],$$

note that we are not taking all the possible permutations as earlier. This is because we may regard f_n as a function $n+1$ variables, where this function is symmetric to its first n variables.

A property of the Skorohod integral is that it is a linear operator. Also for $u \in \text{Dom}(\delta)$ the expectation of the Skorohod integral is zero ($E[\delta(u)] = 0$). This is easily seen from the fact that the Itô integrals have expectation zero. The following theorem tells us under what conditions the Skorohod integral coincides with the Itô integral:

Theorem 2.9. *Let $u = u(t)$, $t \in [0, T]$, be a measurable \mathbb{F} -adapted stochastic process such that*

$$E \left[\int_0^T u^2(t) dt \right] \leq \infty.$$

Then $u \in \text{Dom}(\delta)$ and its Skorohod integral coincides with the Itô integral:

$$\int_0^T u(t) \delta B(t) = \int_0^T u(t) dB(t).$$

Proof. Theorem 2.9 in [4] □

Malliavin Derivative

Originally the Malliavin derivative was constructed on the Wiener space, for this approach see e.g. [4]. We will get an overview of how the Malliavin derivative is defined via chaos expansion. With the knowledge of the chaos expansion we are ready for the definition of the Malliavin derivative for a

stochastic variable $F \in L^2(P)$, where F is \mathcal{F}_T -measurable. As we know from earlier, F then has a chaos expansion and the Malliavin derivative is defined as follows:

Definition 2.10. Let $F \in L^2(P)$ be \mathcal{F}_T -measurable with chaos expansion

$$F = \sum_{n=0}^{\infty} I_n(f_n), \quad \text{where } f_n \in \tilde{L}^2([0, T]^n), n = 1, 2, \dots$$

(i) We say that $F \in \mathbb{D}_{1,2}$ if

$$\|F\|_{\mathbb{D}_{1,2}}^2 := \sum_{n=1}^{\infty} n n! \|f_n\|_{L^2([0, T]^n)}^2 < \infty$$

(ii) If $F \in \mathbb{D}_{1,2}$ we define the Malliavin derivative $D_t F$ of F at time t as the expansion

$$D_t F = \sum_{n=1}^{\infty} n I_{n-1}(f_n(\cdot, t)), \quad t \in [0, T]$$

where $I_{n-1}(f_n(\cdot, t))$ is the $(n-1)$ -fold iterated integral of $f_n(t_1, \dots, t_{n-1}, t)$ with respect to the first $n-1$ variables t_1, \dots, t_{n-1} and $t_n = t$ left as parameter.

Furthermore, there are some computational rules for when one needs to find the Malliavin derivative of a random variable. These rules are also used in the derivation of the Malliavin weight π . The computational rules makes it easier than having to find the chaos expansion, of some stochastic variable F belonging to $L^2(P)$, and applying the definition of the Malliavin derivative.

As in the deterministic case (classic calculus), there is a product rule, a chain rule and a integration by parts formula:

Theorem 2.11 (Product rule). Suppose $F_1, F_2 \in \mathbb{D}_{1,2}^0$. Then $F_1, F_2 \in \mathbb{D}_{1,2}$ and also $F_1 F_2 \in \mathbb{D}_{1,2}$ with

$$D_t(F_1 F_2) = F_1 D_t F_2 + F_2 D_t F_1$$

Proof. Theorem 3.4 in [4] . □

Next we have the chain rule:

Theorem 2.12 (Chain rule). *Let $G \in \mathbb{D}_{1,2}$ and $g \in C^1(\mathbb{R})$ with bounded derivative. Then $g(G) \in \mathbb{D}_{1,2}$ and*

$$D_t g(G) = g'(G) D_t G$$

Here $g'(x) = \frac{d}{dx} g(x)$

Proof. Theorem 3.5 in [4] . □

And at last the integration by parts formula

Theorem 2.13 (Integration by parts). *Let $u(t), t \in [0, T]$, be a Skorohod integrable stochastic process and $F \in \mathbb{D}_{1,2}$, such that the product $Fu(t), t \in [0, T]$, is Skorohod integrable. Then*

$$F \int_0^T u(t) \delta B(t) = \int_0^T F u(t) \delta B(t) + \int_0^T u(t) D_t F dt$$

Proof. Theorem 3.15 in [4] . □

Furthermore the Clark-Ocone formula is a way of representing differentiable stochastic variables via the Malliavin derivative, and takes the form:

Theorem 2.14 (The Clark-Ocone formula). *Let $F \in \mathbb{D}_{1,2}$ be \mathcal{F}_T -measurable. Then*

$$F = E[F] + \int_0^T E[D_t F | \mathcal{F}_t] dB(t)$$

Proof. Theorem 4.1 in [4] . □

One of the reasons why this formula is important in many applications is because the integrand can be expressed explicitly. From the Clark-Ocone formula there is an application to sensitivity analysis, and thus computation of Greeks in finance. We now arrive at the main theorem of this overview of Malliavin calculus, where we need to assume the following in order to obtain the so-called Malliavin weight π :

- We have a general Itô diffusion $X^x(t), t \geq 0$, given by

$$dX^x(t) = b(X^x(t))dt + \sigma(X^x(t))dB(t)$$

where $X^x(0) = x \in \mathbb{R}$, $b : \mathbb{R} \rightarrow \mathbb{R}$ and $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ are given functions in $C^1(\mathbb{R})$ and $\sigma(x) \neq 0$ for all $x \in \mathbb{R}$

- The first variation process $Y(t) := \frac{\partial}{\partial x} X^x(t), t \geq 0$ is given by

$$Y(t) = \exp\left\{\int_0^t \left[b'(X^x(u)) - \frac{1}{2}(\sigma'(X^x(u)))^2\right] du + \int_0^t \sigma'(X^x(u)) dW(u)\right\}$$

- Fixing $T > 0$ and define $g(x) := E^x[\phi(X(T))] = E[\phi(X^x(T))]$

We then have the following theorem:

Theorem 2.15. *Let $a(t)$, $t \in [0, T]$, be a continuous deterministic function such that*

$$\int_0^T a(t) dt = 1.$$

Then

$$g'(x) = E^x \left[\phi(X(T)) \int_0^T a(t) \sigma^{-1}(X(t)) Y(t) dB(t) \right].$$

The random variable

$$\pi^\Delta = \int_0^T a(t) \sigma^{-1}(X(t)) Y(t) dB(t)$$

is a so-called Malliavin weight.

Proof. Theorem 4.14 in [4]. □

Notice that the function $a(t)$ (the weighting function) in the Malliavin weight ($\pi = \pi^\Delta$) is not unique. The weight π allows for a transformation when finding a Greek, which makes it possible to find numerically. The so-called Malliavin weight π , allows for a closed expression of the derivative of a payout process without the derivative of the density function. Hence by this method we are not required to know the density function of the diffusion, nor do we need to demand that the payoff function to be differentiable. However need to know the diffusion.

2.3 Malliavin calculus in case of Lévy processes

In the case of the Brownian motion we now have an overview of concepts in Malliavin calculus, such as the Wiener-Itô chaos expansion, the Malliavin

derivative and the application to sensitivity analysis. This is the reason why we looked into some main concepts of Malliavin calculus. Brownian motion is a continuous stochastic process which does not have jumps. It turns out that Brownian motion is a special case of a more general class of stochastic processes, called Lévy processes, the basics of this will be treated later on. With Lévy processes we have the same concepts, such as the Wiener-Itô chaos expansion, the Skorohod integral etc. However, the notation is more advanced and there are technical differences. We will mention the result, or application due to Malliavin calculus, in sensitivity analysis.

The main difference which concerns the application in sensitivity is the Malliavin derivative. The chain rule in the case of Lévy processes is different, see e.g. Theorem 12.8 in [4], namely that it is a difference operator. In the continuous case we dealt with a differential operator, one can thus not use the same approach to derive the Malliavin weight π .

There are different approaches to derive the Malliavin derivative operator in case of Lévy processes, which does not yield the same operator. One approach could be as a stochastic gradient or through chaos expansion, though they will not yield the same operator, unless we have no jumps, which means we are dealing with a continuous process such as the Brownian motion.

Computation of "Greeks" in the case of jump diffusions

If we are in the pure jump case, we can't use the same approach as we do in the jump diffusion case.¹⁰ This is due to the chain rule for Lévy processes mentioned earlier, one can see that when taking the Malliavin derivative we are actually dealing with a difference operator, more than a differential operator. Hence it will not yield the same result.

A Greek is essentially taking the derivative of a payoff process of the form

$$\frac{\partial}{\partial \theta} E[\phi(S(T))]$$

where θ is a parameter, $\phi(S(T))$ is the payout process of the underlying ($S(T)$ is the underlying asset at time T). This process can be discontinuous, hence it would be hard to obtain the derivative. Next, we have the theorem

¹⁰Note that a pure jump diffusion consists only of jump, while a jump process consists of a continuous part and jump.

which moves the dependence of the derivative away from the payoff function, thus we get a similar Malliavin weight π , as we did in the continuous case. This will help us to solve the problem numerically.

The authors in [4] present the two following Greeks, namely the delta (derivative with respect to the underlying asset) and the gamma (the second derivative with respect to the underlying asset):

Theorem 2.16. *Let $\phi \in L^2(S)$ and let $a \in L^2([0, T])$ be an adapted process such that*

$$\int_0^{t_i} a(t)dt = 1 \quad P - a.e.$$

for all $i = 1, \dots, m$. Then

(1) *The delta of the option is given by*

$$\frac{\partial}{\partial x} E[e^{-rT} \phi(S^x(t_1), \dots, S^x(t_m))] = E[e^{-rT} \phi(S^x(t_1), \dots, S^x(t_m)) \pi^\Delta],$$

where the Malliavin weight π^Δ is given by

$$\pi^\Delta = \int_0^T \frac{a(t)}{x\sigma(t)} dB(t)$$

(2) *The gamma of the option is given by*

$$\frac{\partial^2}{\partial x^2} E[e^{-rT} \phi(S^x(t_1), \dots, S^x(t_m))] = E[e^{-rT} \phi(S^x(t_1), \dots, S^x(t_m)) \pi^\Gamma],$$

where the Malliavin weight π^Γ has the form

$$\pi^\Gamma = (\pi^\Delta)^2 - \frac{1}{x} \pi^\Delta - \frac{1}{x^2} \int_0^T \left(\frac{a(t)}{\sigma(t)} \right)^2 dt.$$

Proof. Theorem 12.29 [4] . □

We observe that in the case of Lévy processes that one does not need to take the derivative of the payoff function. This means we will be able to deal with non-differentiable payoff functions e.g. discontinuous. We also observe that in the setting of jump processes that we are not required to know the explicit density.

Chapter 3

Implementation of Greeks driven by α -stable processes

In the following chapter we will implement a Bismut-Elworthy-Li type formula, which we will often refer to as Zhang's formula [17]. The formula assumes that one has an SDE of the type

$$dX_t(x) = b(t, X_t(x))dt + \sigma dL_t, \quad (3.1)$$

where $b : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a general drift coefficient, which has to be bounded and have continuous first order partial derivatives, and L_t ($0 \leq t \leq T$) is an α -stable process. When we implement the dynamics of the type (3.1), we note that we cannot choose a dynamics where the volatility term, that is the term with the σ , depends on the process $X_t(x)$. We will use an interest rate model for the dynamics (3.1), as we set out to simulate a financial derivative (more specifically a caplet) later on. For the payoff function f , we have to demand that the first order derivative is continuous and bounded, that is $f \in C_b^1(\mathbb{R}^d)$. In order to understand the Bismut-Elworthy-Li-formula of Zhang we need to introduce some background theory. The structure for this chapter is as follows:

- Overview of Lévy processes and how to model Lévy processes (α -stable processes).
- Overview of Zhang's formula, which is a derivative formula of Bismut-Elworthy-Li's type.
- Introduce the Vasicek interest rate model, in the continuous as well as the discontinuous case.

- Implementation of Zhang's formula on a caplet with stochastic interest rate.
- Application to unit-linked policies in life insurance.

The first question that arises when one would like to simulate Lévy processes is how does one build a Lévy process. There are numerous different types of Lévy processes, e.g. Brownian motion, compound Poisson processes. We will use an α -stable process, as this is an assumption in order to implement Zhang's formula.

The α -stable process is a pure jump process, that is, a stochastic process consisting only of jumps. In [17] the stochastic process L_t (in eq. 3.1) is noted on the form $L_t = \{B_{S_t}\}_{t \geq 0}$, this is called a subordinated Brownian motion, which is an α -stable process. When simulating the α -stable process, in the case of Zhang's formula we observe that there is a term with the process S_t , and a process with this process "subscripted" if you may. This technique is known as subordination, that is: we build a Lévy process from a known process, thus we need to look at the concept of subordination. Furthermore we will have a look at the behaviour of the α -stable process, but before this we will deal with the basics of Lévy processes.

We will now look at the more general setting of stochastic processes, namely Lévy processes. This is a class where we allow for stochastic processes with jumps. Why would one be interested in jump processes? In the real world a model which can capture jumps is considered more realistic. There are numerous scenarios in which jumps can occur e.g. an asset can crash overnight if a company goes bankrupt, perhaps there is a sudden change of a bank's monetary policy, which causes the interest rate to jump. The possibilities are endless!

As in the continuous case there are certain properties that a jump process needs to satisfy in order to be a Lévy process. The first part of Lévy processes that is presented, is obtained from first part of chapter 9 in [4]. This framework theory is further on needed for theoretical purposes, such as the proof of the new Bismut-Elworthy-Li-formula in chapter 4. The other part of Lévy processes concerns modeling, in which we will briefly discuss subordination and the α -stable process, this presentation is based on [3]. Readers familiar with the basics of Lévy processes and the α -stable process may skip the first sections, and go straight to section 3.3.

3.1 Lévy processes

We have the following basics of Lévy processes, given a complete probability space (Ω, \mathcal{F}, P) :

Definition 3.1. *A one-dimensional Lévy process is a stochastic process $\eta = \eta(t), t \geq 0$:*

$$\eta(t) = \eta(t, \omega) \quad \omega \in \Omega$$

with the following properties:

- i) $\eta(0) = 0$ *P*-a.s.,
- ii) η has independent increments, that is, for all $t > 0$ and $h > 0$, the increment $\eta(t+h) - \eta(t)$ is independent of $\eta(s)$ for all $s \leq t$,
- iii) η has stationary increments, that is, for all $h > 0$ the increment $\eta(t+h) - \eta(t)$ has the same probability law as $\eta(h)$,
- iv) It is stochastically continuous, that is, for every $t \geq 0$ and $\epsilon > 0$ $\lim_{h \rightarrow 0} P\{|\eta(t+h) - \eta(t)| \geq \epsilon\} = 0$,
- v) η has càdlàg¹ paths, that is, the trajectories are right-continuous with left limits.

A stochastic process η satisfying (i)-(iv) is called a Lévy process in law. When dealing with the continuous case, namely Brownian motion, we see that this is a special case of a Lévy process. It satisfies all of the above properties, but of course need the normality assumption $B(t) \sim N(0, t)$, where $B(t)$ is the Brownian motion. Note that condition four in definition 3.1 says that the probability of observing a jump at a deterministic time point t is equal to zero, meaning that discontinuities only occurs at random times. Hence one does not have sample paths that are continuous.

Let $\eta = \eta(t)$ be a Lévy process, then jump at time t is defined by

$$\Delta\eta(t) := \eta(t) - \eta(t^-)$$

¹The opposite is called càglàd, that is the trajectories are left-continuous with right limits.

If we let $U \in \mathcal{B}(\mathbb{R}_0)$, where $\mathbb{R}_0 := \mathbb{R} \setminus \{0\}$ and $\mathcal{B}(\mathbb{R}_0)$ is the σ -algebra generated by the family of all Borel subsets $U \subset \mathbb{R}$, such that $\bar{U} \subset \mathbb{R}_0$. Then we have that the number of jumps of size $\Delta\eta(s) \in U$ for $0 \leq s \leq t$ is defined by

$$N(t, U) := \sum_{0 \leq s \leq t} \mathbb{1}_U(\Delta\eta(s)), \quad (3.2)$$

where $\mathbb{1}_U$ denotes the indicator function which takes the values 0 or 1. Thus $N(t, U)$ can be regarded as a counting process, where the number of jumps has to be finite, that is $N(t, U) < \infty$, because the paths of η have the càdlàg property. (3.2) defines a Poisson random measure N on $\mathcal{B}(0, \infty) \times \mathcal{B}(\mathbb{R}_0)$ given by

$$(a, b] \times U \longrightarrow N(b, U) - N(a, U), \quad 0 < a \leq b, \quad U \in \mathcal{B}(\mathbb{R}_0).$$

This measure is called the jump measure of η , and its differential form is denoted by $N(dt, dz)$, $t > 0$, $z \in \mathbb{R}_0$. Furthermore we have that the Lévy measure ν of a process η is defined by

$$\nu(U) := E[N(1, U)], \quad U \in \mathcal{B}(\mathbb{R}_0).$$

Note that ν does not necessarily need to be a finite measure, it is possible that

$$\int_{\mathbb{R}_0} \min(1, |z|) \nu(dz) = \infty.$$

This could appear in financial modeling, where it is possible that the trajectories of η could have infinitely many jumps of small sizes.

The Lévy measure always satisfies

$$\int_{\mathbb{R}_0} \min(1, z^2) \nu(dz) < \infty.$$

The measure ν on $\mathcal{B}(\mathbb{R}_0)$ can be a Lévy measure of a Lévy process η if and only if the condition above holds true. This holds true because of the following theorem:

Theorem 3.2 (Lévy-Khintchine formula). *Let η be a Lévy process in law. Then*

$$E[e^{iu\eta(t)}] = e^{i\Psi(u)}, \quad u \in \mathbb{R} \quad (i = \sqrt{-1}),$$

with the characteristic exponent

$$\Psi(u) := i\alpha u - \frac{1}{2}\sigma^2 u^2 + \int_{|z|<1} (e^{iuz} - 1 - iuz)\nu(dz) + \int_{|z|\geq 1} (e^{iuz} - 1)\nu(dz),$$

where the parameters $\alpha \in \mathbb{R}$ and $\sigma^2 \geq 0$ are constants and $\nu = \nu(dz)$, $z \in \mathbb{R}_0$, is a σ -finite measure on \mathbb{R}_0 satisfying

$$\int_{\mathbb{R}_0} \min(1, z^2)\nu(dz) < \infty.$$

It follows that ν is the Lévy measure of η .

Proof. Theorem 9.2 in [4] . □

Combining the jump measure of η , and the Lévy measure ν , we get the so-called compensated jump measure \tilde{N} , (this measure is also known as the compensated Poisson random measure) which is defined as

$$\tilde{N}(dt, dz) := N(dt, dz) - \nu(dz)dt.$$

All Lévy processes can be decomposed, where the idea is to split the continuous part away from the jumps. More precisely we have the following general representation theorem:

Theorem 3.3 (The Lévy-Itô decomposition theorem). *Let η be a Lévy process. Then $\eta = \eta(t)$, $t \geq 0$, admits the following integral representation*

$$\eta(t) = a_1 t + \sigma B(t) + \int_0^t \int_{|z|<1} z \tilde{N}(ds, dz) + \int_0^t \int_{|z|\geq 1} z N(ds, dz) \quad (3.3)$$

for some constants $a_1, \sigma \in \mathbb{R}$. Here $B = B(t)$, $t \geq 0$ ($B(0) = 0$), is a Brownian motion.

Proof. Theorem 9.3 in [4] . □

It's easy to see that the Brownian motion $B(t)$ is a special case of a Lévy process (no jumps will make the integrals in (3.3) become 0). If we assume that the process $\eta(t)$ satisfies

$$E[\eta^2(t)] < \infty \quad t \geq 0, \quad (3.4)$$

then (3.3) takes the form

$$\eta(t) = at + \sigma B(t) + \int_0^t \int_{\mathbb{R}_0} z \tilde{N}(ds, dz). \quad (3.5)$$

Here $a = a_1 + \int_{|z| \geq 1} z \nu(dz)$.

In the framework of Lévy processes, one has a fundamental result, for processes of the type

$$X(t) = x + \int_0^t \alpha(s) ds + \int_0^t \beta(s) dB(s) + \int_0^t \int_{\mathbb{R}_0} \gamma(s, z) \tilde{N}(ds, dz), \quad (3.6)$$

that is, there exists an Itô formula:

Theorem 3.4 (The one-dimensional Itô formula). *Let $X = X(t)$, $t \geq 0$, be the Itô-Lévy process given by (3.6) and let $f : (0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$ be a function in $C^{1,2}((0, \infty) \times \mathbb{R})$ and define*

$$Y(t) := f(t, X(t)), \quad t \geq 0.$$

Then the process $Y = Y(t)$, $t \geq 0$, is also an Itô - Lévy process and its differential form is given by

$$\begin{aligned} dY(t) &= \frac{\partial f}{\partial t}(t, X(t))dt + \frac{\partial f}{\partial x}(t, X(t))\alpha(t)dt + \frac{\partial f}{\partial x}(t, X(t))\beta(t)dB(t) \\ &\quad + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(t, X(t))\beta^2(t)dt \\ &\quad + \int_{\mathbb{R}_0} [f(t, X(t) + \gamma(t, z)) - f(t, X(t)) - \frac{\partial f}{\partial x}(t, X(t))\gamma(t, z)]\nu(dz)dt \\ &\quad + \int_{\mathbb{R}_0} [f(t, X(t^-) + \gamma(t, z)) - f(t, X(t^-))]\tilde{N}(dt, dz). \end{aligned}$$

Proof. Theorem 9.4 in [4]. □

Remark. The Itô formula for the multidimensional case can be found in [4] on page 166.

Looking at (3.5) and setting $\sigma = 0$, we get a so-called pure jump Lévy process. If we let $a = \sigma = 0$ in (3.5), we have that $\eta(t)$ takes the form

$$\eta(t) = \int_0^t \int_{\mathbb{R}_0} z \tilde{N}(ds, dz), \quad t \geq 0.$$

Furthermore, there exists an important representation theorem for stochastic variables in $L^2(P)$ with jumps, namely the Itô representation theorem:

Theorem 3.5. *Let $F \in L^2(P)$ be \mathcal{F}_T -measurable. Then there exists a unique predictable process $\Psi = \Psi(t, z)$, $t \geq 0$, $z \in \mathbb{R}_0$, such that*

$$E \left[\int_0^T \int_{\mathbb{R}_0} \Psi^2(t, z) \nu(dz) dt \right] < \infty$$

for which we have

$$F = E[F] + \int_0^T \int_{\mathbb{R}_0} \Psi(t, z) \tilde{N}(dt, dz)$$

Proof. Theorem 9.10 in [4] . □

3.2 Modeling with the α -stable processes

The α -stable process is a pure jump process, which satisfies the properties of a Lévy process. We would like to model the α -stable process through the concept of subordination. What approach can we take? One option is to construct a Lévy process from known ones. A transformation where the class of Lévy processes is invariant, is through subordination, i.e. through increasing Lévy processes. Subordinators are very important for building Lévy-based models in finance. When using subordination one actually time changes a Lévy process with another increasing Lévy process. An increasing Lévy process has the following properties:

Proposition 3.6. *Let $(X_t)_{t \geq 0}$ be a Lévy process on \mathbb{R} . The following conditions are equivalent:*

- i) $X_t \geq 0$ a.s. for some $t > 0$.
- ii) $X_t \geq 0$ a.s. for every $t > 0$.
- iii) Sample paths of (X_t) are almost surely nondecreasing: $t \geq s \Rightarrow X_t \geq X_s$ a.s.

Proof. See proof of proposition 3.10 in [3]. □

The α -stable processes are frequently used in stochastic modeling. Later on we will encounter and simulate a Bismut-Elworthy-Li's type formula by Zhang [17], which is driven by an α -stable process. To be able to understand this formula it's crucial to understand the driving process, i.e. the α -stable process. First we need the definition of the characteristic function of a random variable X , in order to understand what a stable distribution is:

Definition 3.7. *The characteristic function of the \mathbb{R}^d -valued random variable X is the function $\Phi_X : \mathbb{R}^d \rightarrow \mathbb{R}$ defined by*

$$\forall z \in \mathbb{R}^d \quad \Phi_X(z) = E[\exp(iz \cdot X)] = \int_{\mathbb{R}^d} e^{iz \cdot x} d\mu_X(x),$$

where $\mu_X(x)$ is the density of the random variable X . Furthermore we have the definition of when a random variable $X \in \mathbb{R}^d$ have a stable distribution:

Definition 3.8. *A random variable $X \in \mathbb{R}^d$ is said to have a stable distribution if for every $a > 0$ there exist $b(a) > 0$ and $c(a) \in \mathbb{R}^d$ such that*

$$\Phi_X(z)^a = \Phi_X(zb(a))e^{(ic(a) \cdot z)}, \quad \forall z \in \mathbb{R}^d. \quad (3.7)$$

It is said to have a strictly stable distribution if

$$\Phi_X(z)^a = \Phi_X(zb(a)), \quad \forall z \in \mathbb{R}^d.$$

We have that for every stable distribution there exists a constant $\alpha \in (0, 2]$, which is referred to as the index of stability. In fact such an α will result in the choice $b(a) = a^{\frac{1}{\alpha}}$, which satisfies equation (3.7). To be able to link the stable distributions to Lévy processes, we need to know what lies in the term infinite divisibility, and what the characteristic triplet of a Lévy process is. Every Lévy process have a characteristic triplet ² (A, ν, γ) , which uniquely determines its distribution, where γ is a vector, A is a positive definite matrix and ν is a positive measure. Furthermore we have the definition of infinite divisibility, which is a property where a process can be divided into n independent identically distributed parts:

Definition 3.9. *A probability distribution F on \mathbb{R}^d is said to be infinitely divisible if for any integer $n \geq 2$, there exists n identically independent distributed random variables Y_1, \dots, Y_n such that $Y_1 + \dots + Y_n$ has distribution F .*

Next, we have the fact that any stable distribution is the distribution at a given time of a stable Lévy process:

²For details concerning the characteristic triplet see [3].

Proposition 3.10. *A distribution on \mathbb{R}^d is α -stable with $0 < \alpha < 2$ if and only if it is infinitely divisible with characteristic triplet $(0, \nu, \gamma)$, and there exists a finite measure λ on S , a unit sphere of \mathbb{R}^d , such that*

$$\nu(B) = \int_S \lambda(d\xi) \int_0^\infty 1_B(r\xi) \frac{dr}{r^{1+\alpha}}.$$

A distribution on \mathbb{R}^d is α -stable with $\alpha = 2$ if and only if it is Gaussian.

Proof. Propostion 3.15 in [3]. □

Thus, choosing $\alpha = 2$ yields the Brownian motion.

If we are dealing with a real-valued one-dimensional α -stable distribution ($0 < \alpha < 2$), then the above Lévy measure takes the form

$$\nu(x) = \frac{A}{x^{\alpha+1}} \mathbb{1}_{x>0} + \frac{B}{|x|^{\alpha+1}} \mathbb{1}_{x<0}. \quad (3.8)$$

Let's have a look at some symmetric α -stable distributions.

From figure 3.1 we observe that when α is low ($\alpha = 0.5$), the trajectory is dominated by big jumps. The trajectories resemble a compound Poisson process. On the other hand we observe that when α is big ($\alpha = 1.9$), we have smaller but more frequent jumps, thus we observe that the trajectory resembles a Brownian motion path. Overall we observe that when α increases the jumps get smaller and occur more frequently.

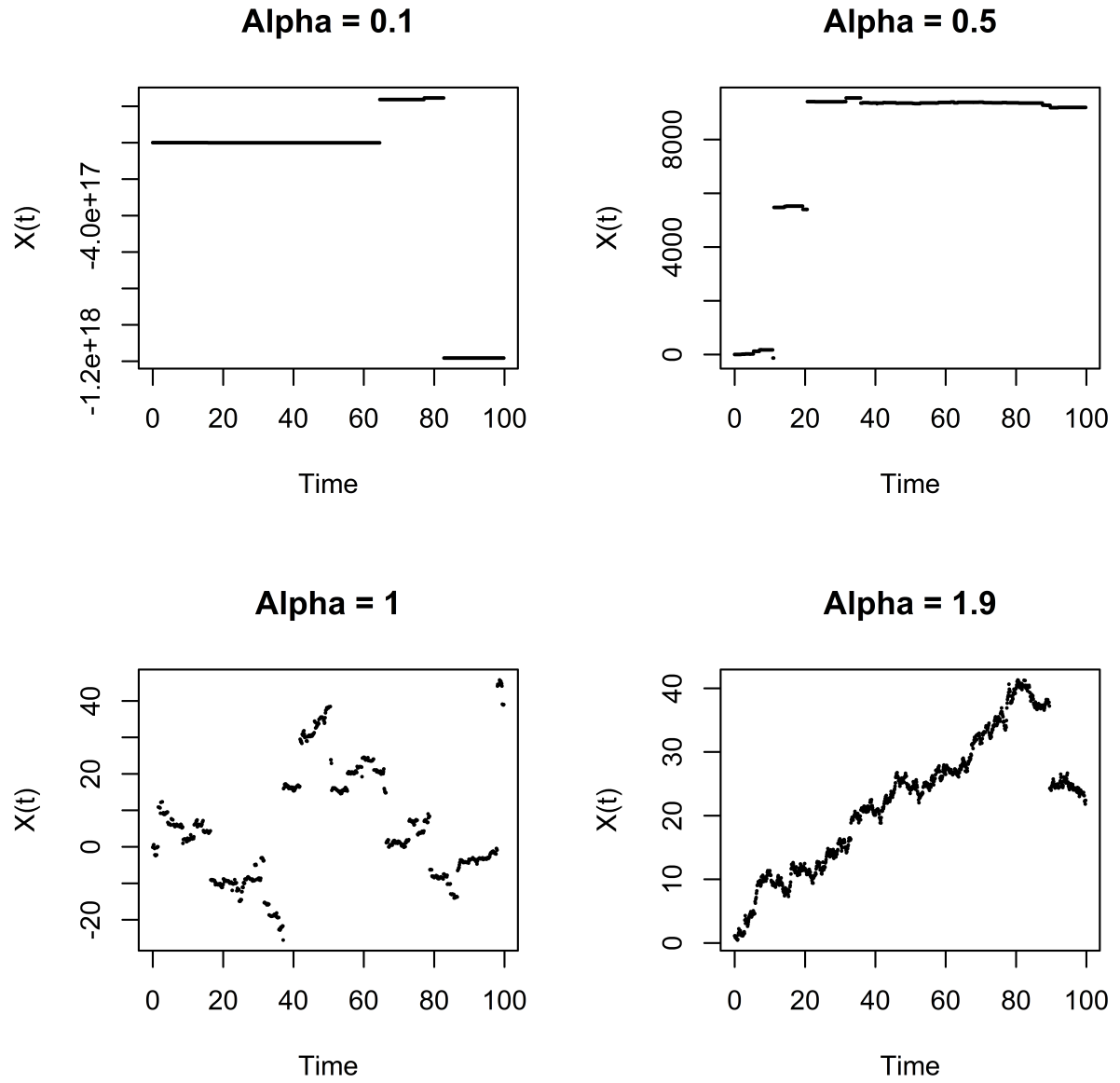


Figure 3.1: α -stable process with different values of α .

3.3 Derivative formula and gradients estimates for SDE's

On the topic of sensitivity analysis, there is a useful tool, or more precisely a type of formula developed by Bismut-Elworthy-Li. This has various applications in functional inequalities, heat kernels estimate and in our topic; sensitivity analysis.

In the paper of Zhang [17] it's proved a derivative formula of Bismut-Elworthy-Li's type for jump diffusion processes. He considers the interesting α -stable process, where he along with the derivative formula derives a gradient estimate for SDEs driven by α -stable noises, where $\alpha \in (0, 2)$. More precisely the formula is for nonlinear SDEs, driven by an α -subordinated Brownian motion (which is an α -stable process). In the literature, work related to this topic has been done by Cass and Fritz (see [17]), where they proved a derivative formula of Bismut-Elworthy-Li's type for SDE's with jumps. Also, Takeuchi (see [17]) worked out a formula for some pure-jump diffusion, with finite moments of all orders. None of the works mentioned above consider the α -stable process, which Zhang considers. Before we proceed to the results of Zhang, let's recall some classical derivative formulas. Let $\{B_t\}_{t \geq 0}$ be a standard d-dimensional Brownian motion. We are considering the following SDE in \mathbb{R}^d :

$$dX_t(x) = b_t(X_t(x))dt + \sigma dB_t, \quad X_0(x) = x, \quad (3.9)$$

where σ is a $d \times d$ invertible matrix, and $b : [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ has continuous first order partial derivatives with respect to x , and

$$\|\nabla b\|_\infty < +\infty,$$

here $\nabla b_s(x) := (\partial_{x^1} b_s(x), \dots, \partial_{x^d} b_s(x))$ and $\|\cdot\|_\infty$ denotes the uniform norm with respect to s and x .

Furthermore there are two forms of derivative formulas: for $f \in C_b^1(\mathbb{R}^d)$ and $h \in \mathbb{R}^d$,

$$\nabla_h E f(X_t(x)) = \frac{1}{t} E \left(f(X_t(x)) \int_0^t \sigma^{-1} [h + (t-s) \nabla_h b_s(X_s(x))] dB_s \right) \quad (3.10)$$

and

$$\nabla_h E f(X_t(x)) = \frac{1}{t} E \left(f(X_t(x)) \int_0^t \sigma^{-1} \nabla_h X_s(x) dB_s \right), \quad (3.11)$$

where for a function φ , $\nabla_h \varphi := \langle \nabla \varphi, h \rangle$ denotes the directional derivative along h . In equation (3.11) ∇b needs to be bounded. $\nabla_h X_t(x)$ satisfies the following linear equation:

$$\nabla_h X_t(x) = h + \int_0^t \nabla b_s(X_s(x)) \cdot \nabla_h X_s(x) ds. \quad (3.12)$$

Let $\alpha \in (0, 2)$ and let $\{S_t\}_{t \geq 0}$ be an independent $\frac{\alpha}{2}$ -stable subordinator, that is, an increasing \mathbb{R} -valued process with stationary independent increments, and

$$E[e^{iuS_t}] = e^{t|u|^{\alpha/2}}, \quad u \in \mathbb{R} \quad (i = \sqrt{-1}).$$

Letting S_t be defined in this way we have that the subordinated Brownian motion $\{B_{S_t}\}_{t \geq 0}$ is an α -stable process.

Zhang's formula assumes an SDE in \mathbb{R}^d driven by B_{S_t} :

$$dX_t(x) = b_t(X_t(x))dt + \sigma \cdot dB_{S_t}, \quad X_0(x) = x.$$

Then we have the following theorem and what we refer to as Zhang's formula, that is, a formula for $\nabla E f(X_t(x))$:

Theorem 3.11. *Under the condition that $b : [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ has continuous first order partial derivatives with respect to x and $\|\nabla b\|_\infty < +\infty$, for any function $f \in C_b^1(\mathbb{R}^d)$ and $h \in \mathbb{R}^d$, we have*

$$\nabla_h E f(X_t(x)) = E \left(\frac{1}{S_t} f(X_t(x)) \int_0^t \sigma^{-1} \cdot \nabla_h X_s(x) dB_{S_s} \right), \quad (3.13)$$

where $\nabla_h X_s(x)$ is determined by equation (3.12). In particular, for any $\alpha \in (0, 2)$ and $p \in (1, \infty]$, there exists a constant $C = C(\alpha, p) > 0$ such that for all $t > 0$,

$$|\nabla E f(X_t(x))| \leq C \|\sigma^{-1}\| e^{\|\nabla b\|_\infty t} t^{-\frac{1}{\alpha}} (E|f(X_t(x))|^p)^{\frac{1}{p}}, \quad (3.14)$$

where $\|\sigma^{-1}\| := \sup_{|x|=1} |\sigma^{-1}x|$ and $|\cdot|$ denotes the Euclidean norm.

Remark. By equation (3.12) we see that $s \mapsto \nabla_h X_s(x)$ is a bounded and continuous $\sigma\{B_{S_r} : r \leq s\}$ -adapted process. So the stochastic integral in (3.13) makes sense.

Proof. See Theorem 1.1 in [17]. □

3.4 Simulation of the derivative of a caplet with respect to the initial interest rate, under the Vasicek interest rate model

In this section we will run simulations of eq. (3.13), with a stochastic interest rate (Vasicek model), with the α -stable process, using high values of α , that is α close to two. Then we will observe what the behavior is like when applying the Brownian motion to Zhang's formula, since we have the fact that an α equal to two yields the Brownian motion.

3.4.1 The Vasicek interest rate model

The Vasicek model was introduced by Oldřich Vašíček in 1977, and was one of the first interest rate models to capture mean reversion. Mean reversion is the effect that either a high or a low interest rate will tend back to its average. The average could for instance be determined by a country's monetary policy. In the real world a justification for using such a model, i.e. capture the mean reversion effect, is because of too high interest rates will cause the economy tend to slow down. A reason for this effect is that it becomes less profitable to invest (e.g. borrow money from the bank). Hence an economy running slowly would result in a lower interest rate. On the other hand, if the interest rate is too low its not healthy for the economy, e.g. this could result in foreign individuals from an arbitrary country, not demanding their currency because it's not profitable (to e.g. save money in banks). This will in turn result in a bad currency rate, which is not healthy for the economy over time. Thus, the interest rate would be driven up and tend to its average.

In the setting of Brownian motion the Vasicek model assumes that the short rate evolves as an Ornstein-Uhlenbeck process, which is the solution of an SDE on the form:

$$dX_t = \mu X_t dt + \sigma dB_t.$$

More precisely the Vasicek short rate model is given by the following SDE:

$$dr_t = a(b - r_t)dt + \sigma dB_t, \quad r(0) = r_0, \quad (3.15)$$

which has the solution (for proof see Lemma A.3.)

$$r_t = r_0 e^{-at} + b(1 - e^{-at}) + \sigma e^{-at} \int_0^t e^{as} dB_s.$$

In the case of a jump diffusion we have that (L_s denotes a Lévy process):

$$r_t = x + \int_0^t a(b - r_s)ds + \int_0^t \sigma dL_s, \quad r_0 = x \quad (3.16)$$

The parameters r_0, a, b and σ are non-negative, and can be interpreted as follows:

- r_0 - Initial interest rate.
- a - The speed at which the trajectories will go towards the mean (b) in time.
- b - Mean long term interest rate level.
- σ - The volatility.

Dealing with Brownian motion we see that r_t is normally distributed with mean

$$E[r_t] = r_0 e^{-at} + b(1 - e^{-at}),$$

by using the fact that the expectation of an Itô integral is zero. Furthermore r_t has variance

$$\begin{aligned} \text{Var}[r_t] &= \text{Var} \left[\sigma e^{-at} \int_0^t e^{as} dB_s \right] \\ &= \sigma^2 e^{-2at} \int_0^t (e^{as})^2 ds \\ &= \frac{\sigma^2}{2a} (1 - e^{-2at}), \end{aligned} \quad (3.17)$$

where the second equality follows from the Itô isometry. If we look at the average overnight interest rate, namely $E[r_t]$ and let $t \rightarrow \infty$ we see that

$$\lim_{t \rightarrow \infty} E[r_t] = b \quad \text{and} \quad \begin{cases} E[r_t] \nearrow b & \text{if } b \geq r_0 \\ E[r_t] \searrow b & \text{if } b < r_0. \end{cases}$$

One deficiency of the Vasicek model is that the interest rate r_t , can become negative, and the stochastic noise term does not depend on the evolution of r_t . On the other hand it is not a too complex model, in the sense that

it can be solved straight forward, in the setting of Brownian motion and it is possible to estimate the parameters to historical data, by for instance maximum likelihood estimation. One of the main reasons why we choose the Vasicek model is because when we simulate equation (3.13) we have to use a model satisfying equation (3.9), which demands that the volatility term does not depend on the process itself. One might argue that this a limitation of the forthcoming simulations, as one would expect that an interest rate does not have constant volatility, but should depend on the interest rate level. The Black-Karasinsky is a model which captures this effect, but cannot be solved analytically as the Vasicek can. An alternative model could for instance be the Cox Ingersoll Ross (CIR) model, which have advantages and disadvantages compared to the Vasicek model. For more on this and other interesting interest rate models see [2].

In figure 3.2 we see a simulation of Vasicek with Brownian motion (drift= 0 and volatility = 0.1), where the black line indicates the expectation, which we see is b in the long run. The oscillating graphs around the mean, with colors, are the result of a run with one omega (one path).

Vasicek with $a = 0.05$, $b = 0.04$, $r_0 = 0.044$ and sigma = 0.004

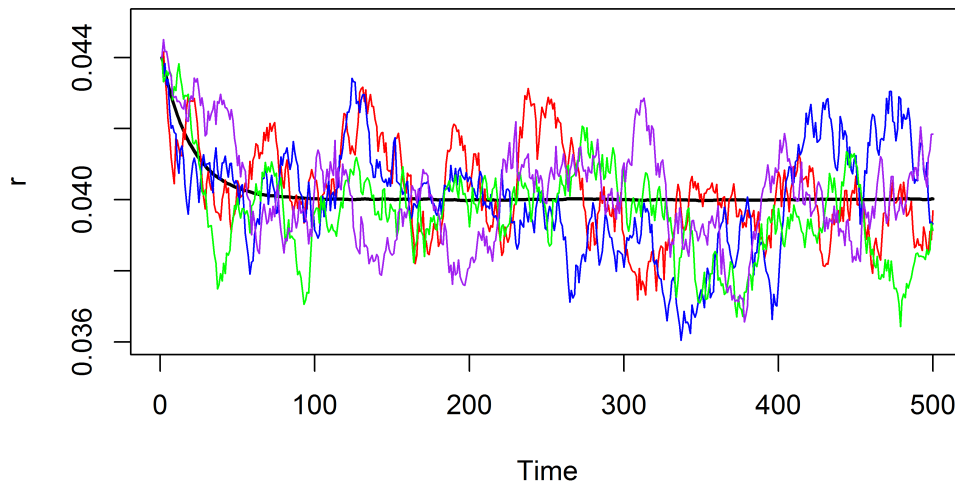


Figure 3.2: Simulation of the Vasicek model with Brownian motion.

A simulation is also done for the Vasicek model in the case of a pure jump process, namely the α -stable process. This can be seen in figure 3.3, where the black line is the average, which is obtained by simulating ten thousand paths of the Vasicek model, and taking the average at each time point. The estimate barley fluctuates more than in the case of Brownian motion. The red line is one path, where we observe jumps in the interest rate and the property of mean reversion effect, i.e. when the interest is far away from the longterm level it will be pulled back. Note how large the jumps are for this (red) path. They might be a little big, compared to what one could expect from the real world.

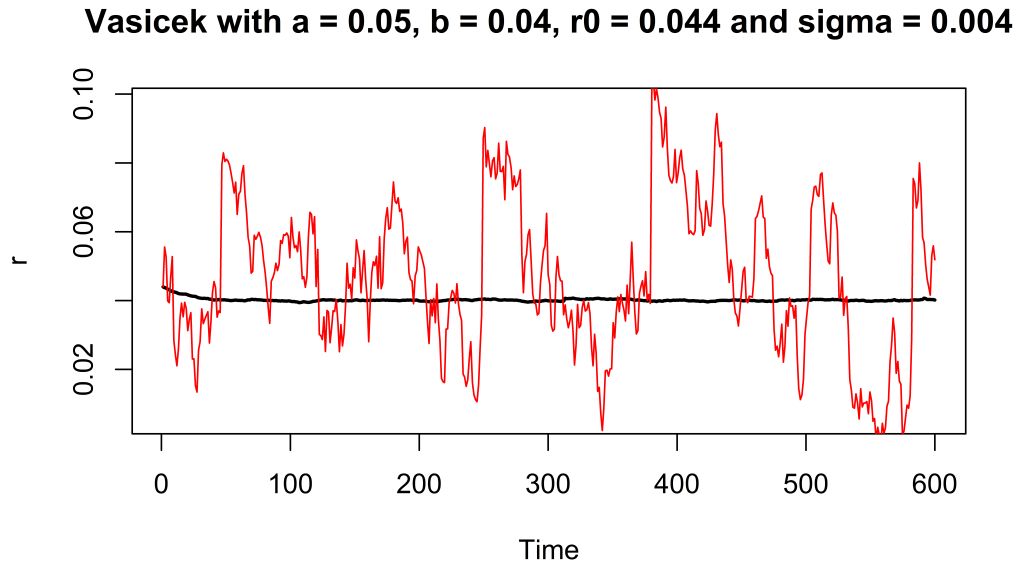


Figure 3.3: Simulation of the Vasicek model with an α -stable process.

3.4.2 Simulation of the derivative of a caplet with respect to the initial interest rate

Applying the Vasicek model to eq. (3.13), we will look at a caplet ³, i.e. $\max(r_t - K, 0)$, where r_t is the interest rate and K is the strike price. This

³One could think of a caplet as a call option, with underlying being the interest rate.

could e.g. be the lowest interest rate an investor can handle. Thus he would like to buy insurance to avoid a too low interest rate ⁴. Then we will take the derivative of the caplet with respect to the initial interest $r_0 = x$, and see how this process evolves. The underlying process will be an α -stable process, that is, the B_{S_t} -term will be an α -stable process. The $\{S_t\}_{t \geq 0}$ -term denotes the $\frac{\alpha}{2}$ subordinator. Applying eq. (3.13) we have the following:

$$f(X_t(x)) = \max(r_t - K, 0),$$

where r_t is determined by the Vasicek interest rate model and K denotes the strike price. Recall that in the scenario of jumps, the Vasicek model has the following form:

$$r_t = x + \int_0^t a(b - r_s)ds + \int_0^t \sigma dL_s, \quad (3.18)$$

where L_s is a Lévy process. Taking the derivative of eq. (3.18) with respect to the initial interest rate x , yields

$$\frac{d}{dx}r_t = 1 + \int_0^t -a \frac{d}{dx}r_s ds,$$

this a deterministic differential equation in which we get the solution $\frac{d}{dx}r_t = e^{-at}$, hence $r_t = xe^{-at}$. Arriving at σ , the $d \times d$ invertible matrix, we observe that in our case that we have a 1×1 matrix, thus the inverse σ^{-1} is just $\frac{1}{\sigma}$. Inserting these ingredients into eq. (3.13) we get the following:

$$\begin{aligned} \frac{d}{dx}E[\max(r_t(x) - K, 0)] &= E \left[\frac{1}{S_t} \max(r_t(x) - K, 0) \int_0^t \frac{1}{\sigma} \frac{d}{dx}r_s(x) dB_{S_s} \right] \\ &= E \left[\frac{1}{S_t} \max(r_t(x) - K, 0) \int_0^t \frac{1}{\sigma} \left(1 + \int_0^s -a \frac{d}{dx}r_u(x) du \right) dB_{S_s} \right] \\ &= E \left[\frac{1}{S_t} \max(r_t(x) - K, 0) \int_0^t \frac{1}{\sigma} e^{-as} dB_{S_s} \right]. \quad (3.19) \end{aligned}$$

We will simulate eq.(3.19), using the Vasicek model where time goes from 0 to 600, in say days, where we observe a daily update on the interest rate, meaning we use $n = 600$ uniformly distributed points in the Vasicek model. For our underlying process, namely the α -stable process and $\frac{\alpha}{2}$ -subordinator,

⁴If the investor can't handle a high interest rate, then the investor can purchase the opposite product, namely a floorlet.

we will choose a high α (meaning close to 2), say $\alpha = 1.999$. Then we will compare this to the case when $\alpha = 2$, in which we are in the setting of the Brownian motion. In the scenario $\alpha = 2$ we let $B_{S_t} = B_t$, where $B_t \sim N(0, t)$ is the Brownian motion, thus S_t becomes t under the expectation. Inserting these ingredients into eq. (3.19) yields

$$\frac{d}{dx} E [\max(r_t(x) - K, 0)] = E \left[\frac{1}{t} \max(r_t(x) - K, 0) \int_0^t \frac{1}{\sigma} e^{-as} dB_s \right]. \quad (3.20)$$

Simulating eq.(3.19) $m = 1,000,000$ times and obtaining the average by the weighted average of each time point, we estimate

$$E(X_{t_i}) \approx \frac{1}{m} \sum_{j=1}^m \frac{1}{\{S_{t_i}\}_j} \max(r_{t_i}(x) - K, 0) \frac{1}{\sigma} \int_0^{t_i} e^{-as} d\{B_{S_s}\}_j.$$

Note that S_{t_i} and B_{S_s} are independent of each other for each j .

Implementing this in R, using a caplet on the interest rate, i.e. $\max(r_t - K, 0)$. Where the strike price is set to $K = 0.05$, $r_0 = 0.048$, long term interest rate $b = 0.04$, $a = 0.05$ (rate of which interest rate tends towards b), the volatility $\sigma = 0.004$, which is used on both $\alpha = 1.999$ and $\alpha = 2$, where we know the latter value for α means we are in the setting of Brownian motion. When simulating the Brownian motion we arbitrary choose drift = 0.06 and volatility = 0.5. These parameters could e.g. be used by an investor who borrows money from a bank, to invest and needs insurance in case of a high interest rate, here 5%, which might be due to the costs related to paying interest on the loan the investor took out. The investor knows the longterm interest rate is 0.04 (by the country's monetary policy), but also knows the rate fluctuates and hence wants protection. How sensitive is this with respect to the initial interest rate?

$r_0 = 0.048$, strike = 0.05, longterm interest rate = 0.04

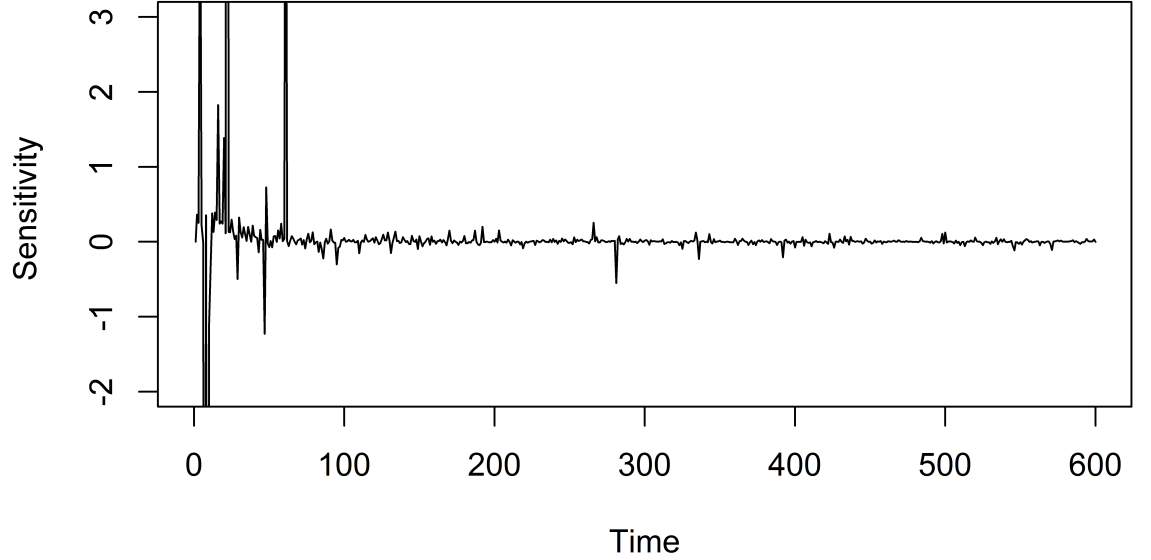


Figure 3.4: Simulation of eq.(3.19).

We see in figure 3.4 that the initial interest rate is very sensitive in the beginning and tend to oscillate around zero, where we see the oscillations decrease as time passes by. The intuitive reason is that the sensitivity of the caplet with respect to the initial interest rate, should decrease (i.e. tend to zero) as time passes by. This is because of the fact that the interest rate years back will have a lower impact on today's value on the interest rate, as time passes by. Note that the random spikes after a period of time (say after day 250) could be caused by jumps in the interest rate or random error, due to the behavior of the α -stable process. If we zoom in on figure 3.4 from 0 to 400 days and add a simulation of Brownian motion (red line) on top, we observe that it tends fast to zero. One reason is because the $\frac{1}{t}$ term in eq. (3.20) will dominate as the time passes by. It is quite interesting how a high α behaves almost the same way as the Brownian motion, in which $\alpha = 2$. However the α -stable process is more realistic in the real market, as it can capture jumps. We observe the difference between the jump process

and continuous process, by small oscillations around zero. This gives proof of the importance of sensitivity analysis in the setting of a process with jumps⁵ and not just a continuous process.

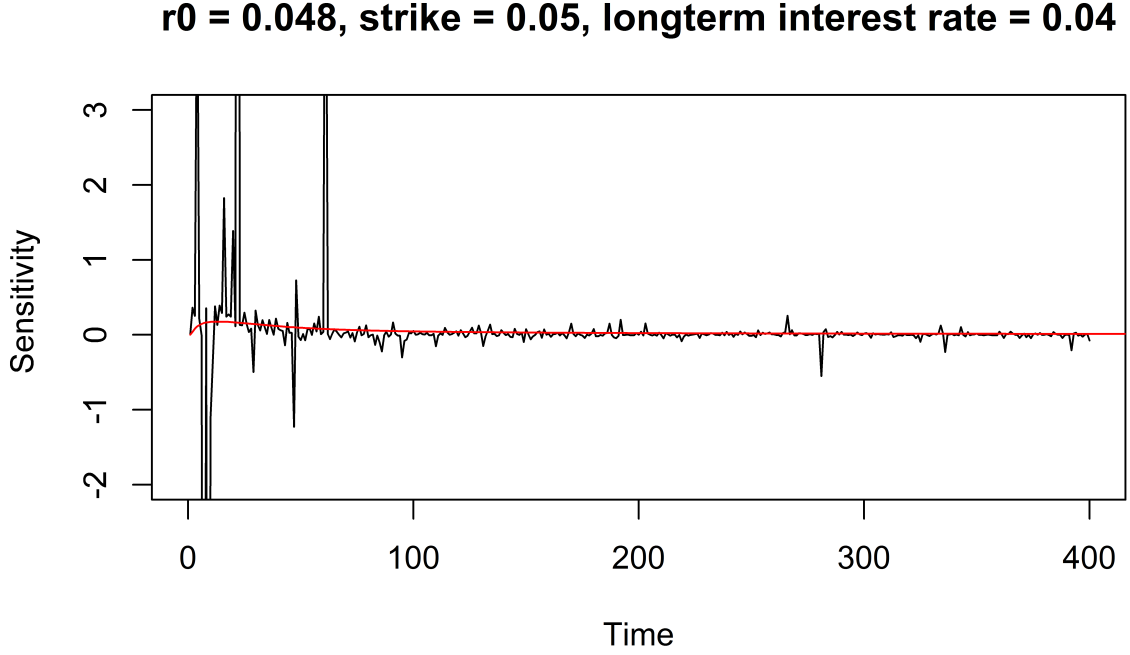


Figure 3.5: Simulation of eq.(3.19) with $\alpha = 1.999$ and $\alpha = 2$ (red).

3.5 Application to life insurance

In the following section we will consider an application of Zhang's formula to life insurance products. We will utilize the Bismut-Elworthy-Li type formula from the previous section, to simulate the sensitivities of unit-linked policies. First we will introduce the framework of life insurance mathematics we will work in. As a brief introduction we have some examples from classical life insurance contracts:

⁵Recall that the α -stable process is a pure jump process, and not a general jump process with a continuous term.

- **Pure endowment:** Payment from the insurer to the insured, when the insured reaches ages of maturity of the policy. In case of death before reaching the time of maturity of the policy, there is no payment.
- **Term-life insurance:** If the individual dies before the age of maturity of the policy, the heirs receive a payment.
- **Endowment:** Sum of a pure endowment insurance and a term life insurance. Which means that the individual holding such a contract will receive a payment (to the heirs), in case of an early death. If the individual reaches the time of maturity of the policy, then the individual receives payment.

The following presentation of the framework in life insurance, is based on chapter 2 in [11]. In the following section we will present an application to life-insurance products, more precisely unit linked policies.

3.5.1 Framework in life insurance

To be able to deal with life insurance we need some basic knowledge of transition probabilities between states, e.g. probability of going from the state alive to dead or alive to disabled. For our purposes we will only need the state alive or dead, denoted respectively by $(*, \dagger)$. To obtain the transition probabilities we will resort to Kolmogorov's differential equations, where we first need the knowledge of what transition rates are, and what a regular Markov chain is.

Definition 3.12. Let $(X_t)_{t \in T}$ be a Markov chain with finite state space S and $T \subset \mathbb{R}$. For $N \subset S$ we define

$$p_{jN}(s, t) := \sum_{k \in N} p_{jk}(s, t).$$

Definition 3.13. (Transition rates) Let $X(t)_{t \in T}$ be a Markov chain in continuous time with finite state space S . $(X_t)_{t \in T}$ is called regular, if

$$\mu_i(t) = \lim_{\Delta t \searrow 0} \frac{1 - p_{ii}(t, t + \Delta t)}{\Delta t} \quad \text{for all } i \in S \quad (3.21)$$

$$\mu_{ij}(t) = \lim_{\Delta t \searrow 0} \frac{p_{ij}(t, t + \Delta t)}{\Delta t} \quad \text{for all } i \neq j \in S \quad (3.22)$$

are well defined and continuous with respect to t .

The functions $\mu_i(t)$ and μ_{ij} are called transition rates. Also $\mu_{ii}(t) \stackrel{\text{def}}{=} -\mu_i(t)$, $i \in S$.

Remark. • The transition rates can be interpreted as derivatives of the transition probabilities. For $i \neq j$ we get

$$\begin{aligned}\mu_{ij}(t) &= \lim_{\Delta t \searrow 0} \frac{p_{ij}(t, t + \Delta t) - p_{ij}(t, t)}{\Delta t} \\ &= \lim_{\Delta t \searrow 0} \frac{p_{ij}(t, t + \Delta t) - p_{ij}(t, t)}{\Delta t} \\ &= \frac{d}{ds} p_{ij}(t, s)|_{s=t}.\end{aligned}$$

- $\mu_{ij}dt$ can be interpreted as the infinitesimal transition rate from i to j in the time interval $[t, t + dt]$. Furthermore $\Lambda(t)$ is defined as

$$\Lambda(t) = \begin{bmatrix} \mu_{11}(t) & \mu_{12}(t) & \mu_{13}(t) & \dots & \mu_{1n}(t) \\ \mu_{21}(t) & \mu_{22}(t) & \mu_{23}(t) & \dots & \mu_{2n}(t) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{n1}(t) & \mu_{n2}(t) & \mu_{n3}(t) & \dots & \mu_{nn}(t) \end{bmatrix}$$

where Λ generates the behavior of a Markov chain. Hence we have that

$$\Lambda = \lim_{\Delta t \rightarrow 0} \frac{P(\Delta t) - 1}{\Delta t}.$$

$\Lambda := \Lambda(0)$ is called the generator of the one parameter semigroup. $P(t)$ can be reconstructed by

$$P(t) = \exp(t\Lambda) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \Lambda^n$$

How do we obtain the transition probabilities from one state to another? Given a state space, consisting of different states e.g. alive, ill, critically ill and dead, how does one find the probabilities? The answer, is to solve the Kolmogorov differential equations:

Theorem 3.14. (*Kolmogorov*) *Let $(X_t)_{t \in T}$ be a regular Markov chain on a finite state space S . Then the following statements hold:*

1. (Backward differential equations)

$$\frac{d}{ds}p_{ij}(s, t) = \mu_i(s)p_{ij}(s, t) - \sum_{k \neq i} \mu_{ik}(s)p_{kj}(s, t), \quad (3.23)$$

$$\frac{d}{ds}P(s, t) = -\Lambda(s)P(s, t) \quad (3.24)$$

2. (Forward differential equations)

$$\frac{d}{dt}p_{ij}(s, t) = -p_{ij}(s, t)\mu_j(t) + \sum_{k \neq j} p_{ik}(s, t)\mu_{kj}(t), \quad (3.25)$$

$$\frac{d}{dt}P(s, t) = P(s, t)\Lambda(t). \quad (3.26)$$

Proof. See proof of Theorem 2.3.4 in [11]. \square

As mentioned earlier, we will need the transition probabilities for a model concerning the two states alive and dead (*, †). The simplest transition rate of $\mu_{*\dagger}$ would be to set it equal to a constant. This is not very realistic, as one can imagine, there is a higher probability of not dying given that the person is young (say 20 years old), than an old person (say 80 years old). We will use a death rate estimated from [11]. More precisely a polynomial of degree two was fitted to $\log(\mu_{*\dagger})$:

$$\mu_{*\dagger} = \exp(-7.85785 + 0.01538 \cdot x + 5.77355 \cdot 10^{-4} \cdot x^2) \quad (3.27)$$

Using the fact that

$$\mu_i(t) = \sum_{j \neq i} \mu_{ij}(t)$$

and the Kolmogorov forward differential equations (eq. 3.25) we get

$$\frac{d}{dt}p_{**}(s, t) = -p_{**}(s, t)\mu_*(t) + p_{*\dagger}(s, t)\mu_{\dagger*}(t) \quad (3.28)$$

$$\frac{d}{dt}p_{*\dagger}(s, t) = -p_{*\dagger}(s, t)\mu_{\dagger}(t) + p_{**}(s, t)\mu_{*\dagger}(t), \quad (3.29)$$

where we note that $\mu_{\dagger} = 0$ and $\mu_{\dagger*} = 0$. Solving eq. (3.28) yields

$$\begin{aligned}
\frac{d}{dt}p_{**}(s, t) &= -p_{**}(s, t)\mu_{*\dagger}(t) \\
\frac{\frac{d}{dt}p_{**}(s, t)}{p_{**}(s, t)} &= -\mu_{*\dagger}(t) \\
\frac{d}{dt}\ln(p_{**}(s, t)) &= -\mu_{*\dagger}(t) \\
p_{**}(s, t) &= \exp\left(-\int_s^t \mu_{*\dagger}(u)du\right). \tag{3.30}
\end{aligned}$$

As for eq. (3.29) we note that since there are two states, namely dead and alive that we only have the transition probabilities $p_{**}(s, t)$ and $p_{*\dagger}(s, t)$ (And of course the trivial states $p_{\dagger\dagger}(s, t) = 1$ and $p_{\dagger*}(s, t) = 0$). Thus the solution for $p_{*\dagger}(s, t)$ becomes

$$p_{*\dagger}(s, t) = 1 - \exp\left(-\int_s^t \mu_{*\dagger}(u)du\right) \tag{3.31}$$

3.5.2 Unit-linked policies

In the world of life insurance there are numerous contracts and policies, e.g. a lump sum payment in case of death, disability payments if one becomes ill, pension schemes etc. Unit-linked policies are insurance contracts where the policyholder gets the benefit of both investment and insurance. The investments could for instance be in stocks or bonds. This type of policy usually has a single premium. We will look at two unit-linked policies, namely an endowment policy, which is represented in the following way:

$$V(0) = E_Q[\exp(-\delta T)C(T)] \cdot {}_TP_s. \tag{3.32}$$

Here we have that

$${}_TP_s = p_{**}(s, s+1) \cdot p_{**}(s+1, s+2) \cdots p_{**}(s+T-1, s+T)$$

denotes the probability of surviving over the period of the contract, when the individual is s years old at the beginning of the contract, and T is number of periods (years) ahead.

Furthermore we have a term life insurance:

$$V(0) = \int_0^T E_Q[\exp(-\delta t)C(t)]p_{**}(s, s+t)\mu_{*\dagger}(s+t)dt. \tag{3.33}$$

Here $V(0)$ denotes the single premium. The expectation is taken with respect to the risk neutral measure Q , $\exp(-\delta t)$ is the discount factor where δ is a constant. $C(t)$ is the payoff function (where $C(t)$ is stochastic), which could e.g. be $\max(r_t - K, 0)$ (caplet) where r_t is a stochastic interest rate and K is the strike price. $C(t)$ could also e.g. be a guaranteed return, i.e. $\max(N(t)S(t), G(t))$ where $N(t)$ is the number of shares bought of an arbitrary stock. Where $S(t)$ denotes the stochastic stock price function, e.g. $S(t) = xe^{L_t}$ for a Lévy process L_t , the initial value x of the stock and $G(t)$ the guarantee are set equal to a fixed number.

One might be interested in what $V(t)$ is in general, not just $V(0)$. If we assume we have a filtration \mathcal{F}_t of information up to time t , which is a σ -algebra, then we can have a look at the sensitivity of $V(t)$. Let us consider the sensitivity of an endowment policy with respect to an x , where x is a parameter of the stochastic payout function $C(t)$. Using the Markov property of r_t (we assume that r_t is the risk neutral process), we have the following for $V(t)$, at a future time point t .

$$\begin{aligned}\frac{d}{dx}V(t) &= \frac{d}{dx}E_Q[e^{-\delta T}\max(r_T - K, 0)|\mathcal{F}_t] \cdot {}_TP_s \\ &= e^{-\delta T}\frac{d}{dy}E_Q[\max(r_{T-t}^{y,t} - K, 0)]|_{y=r_t^x} \cdot \frac{d}{dx}r_t^x \cdot {}_TP_s\end{aligned}\quad (3.34)$$

Similar computations for the term life insurance, with say a put option with strike price K (makes the payoff function $C(t)$ bounded). Where the number of shares bought are noted by $N(t)$ and the guarantee is noted by $G(t)$. Assuming both of the latter functions are constants, we obtain:

$$\begin{aligned}\frac{d}{dx}V(t) &= \frac{d}{dx}\int_t^T E_Q[e^{-\delta u}\max(K - N(u)S(u), G(u))|\mathcal{F}_t] \cdot p_{**}(s, s+u)\mu_{*\dagger}(s+u)du \\ &= \int_t^T e^{-\delta u}\frac{d}{dx}E_Q[\max(K - N(u)S(u), G(u))|\mathcal{F}_t] \cdot p_{**}(s, s+u)\mu_{*\dagger}(s+u)du \\ &= \int_t^T e^{-\delta u}\frac{d}{dy}E_Q[\max(K - N(u)S^{y,t}(u-t), G(u))]|_{y=S(t)} \\ &\quad \cdot \frac{d}{dy}S^y(t) \cdot p_{**}(s, s+u)\mu_{*\dagger}(s+u)du\end{aligned}\quad (3.35)$$

Where $S(t)$ could e.g. be an exponential Lévy process, i.e.

$$S(t) = xe^{\sigma L_t}, \quad \text{where } x = S(0),$$

so

$$\begin{aligned}\frac{d}{dx}S(t) &= \frac{d}{dx}xe^{\sigma L_t} \\ &= e^{\sigma L_t}.\end{aligned}\tag{3.36}$$

Using eq. (3.36) in connection with the sensitivity of the term life insurance (eq. (3.35)) we obtain

$$\begin{aligned}\frac{d}{dx}V(t) &= \int_t^T e^{-\delta u} \frac{d}{dy} E_Q[\max(K - N(u)S(u-t), G(u))]|_{y=S(t)} \\ &\quad \frac{d}{dy}S(t) \cdot p_{**}(s, s+u)\mu_{*\dagger}(s+u)du \\ &= \int_t^T e^{-\delta u} \frac{d}{dx} E_Q[\max(K - N(u)ye^{\sigma L_{u-t}}, G(u))]|_{y=S(t)} \cdot e^{\sigma L_t} \\ &\quad p_{**}(s, s+u)\mu_{*\dagger}(s+u)du\end{aligned}\tag{3.37}$$

Furthermore in both eq. (3.34) and eq. (3.37) we can apply Zhang's formula, namely eq. (3.13) for the derivative of the expectation to obtain a solution.

3.5.3 Simulation of the sensitivity of unit-linked policies

Let's have a look at eq. (3.34) at $t = 0$ (beginning of the contract) in connection with a caplet, i.e. let $C(t) = \max(r_t - K, 0)$ where we let the interest rate process r_t be modeled by the Vasicek model with jumps (see eq. (3.18)). The parameters for the Vasicek model are chosen to be

- $r_0 = 0.048$ (Initial interest rate).
- $a = 0.05$ (Rate of convergence to the longterm interest rate).
- $b = 0.04$ (Longterm interest rate).
- $\sigma = 0.004$ (Volatility).

Sensitivity of an endowment policy

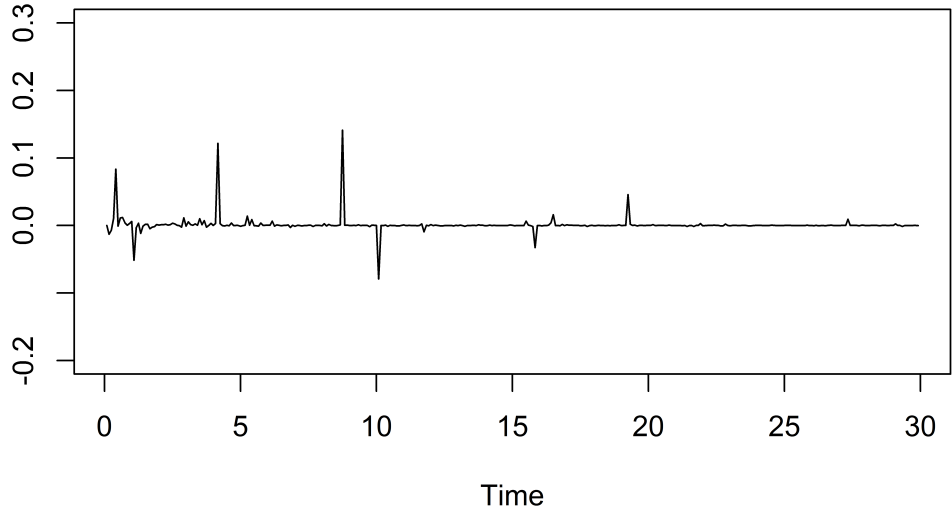


Figure 3.6: Simulation of eq.(3.34) with $\alpha = 1.9$.

The strike K is set to 0.05, and the fixed interest rate for the discount factor is $\delta = 0.03$. We let the contract run for an individual aged 35 until age 65, with the probability of dying is given by eq.(3.27) and eq.(3.31). Simulating the expectation 200,000 times, with the symmetric α stable process for $\alpha = 1.9$ yields figure 3.6. We observe that the sensitivity of the endowment policy decreases. At earlier time points there are peaks, which might be due to the α -stable process or perhaps jumps of great magnitude, which blows up the effect of the initial interest rate. As time goes by we observe very small oscillations around zero, when look from time point say twelve and out. Which is what one would expect, namely that the initial interest has a smaller effect as the time increases. Recall the simulation of eq.(3.19), where we observed that the expectation tended towards zero as time went on, we would also expect this from the endowment. This is due to the fact that the expectation, multiplied by the probability of staying alive during the contract, (which is between zero and one) and the fact that the discount factor is small, will make the expression even smaller.

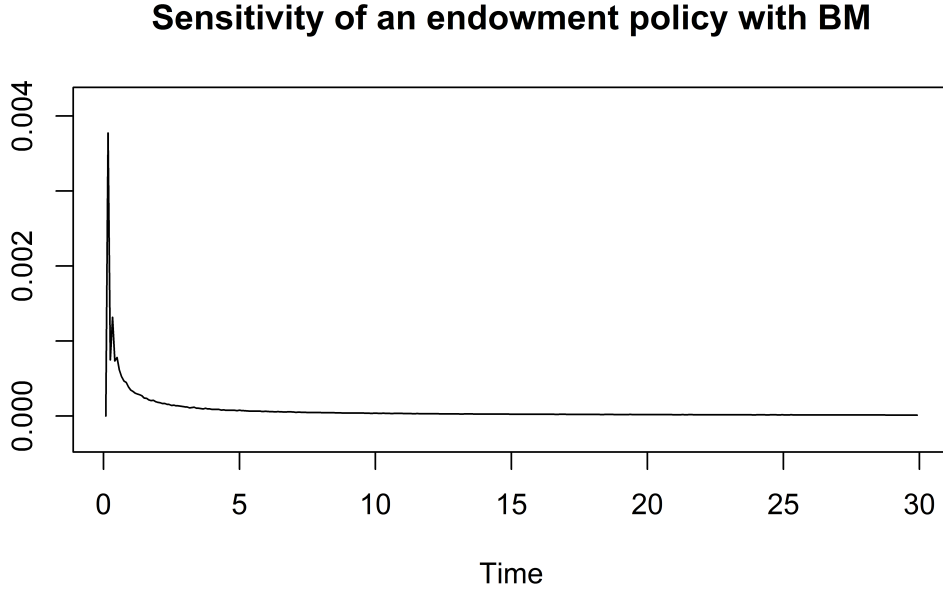


Figure 3.7: Simulation of eq.(3.34) with Brownian motion.

We chose a high value for α and we know that $\alpha = 2$ makes our process become the Brownian motion. So what does the sensitivity of the same endowment policy become, with respect to the initial interest rate, when we are dealing with α equal to two? If we let the parameters of the Brownian motion take the arbitrary values $\mu = 0.06$ (drift) and $\sigma = 0.5$ (volatility), we see by figure 3.7 that the sensitivity tends fast towards zero. The magnitude of the sensitivity is a lot smaller in the Brownian motion case, than in the setting of the α -stable process.

Conducting simulations on the sensitivity of the term life insurance, namely eq. (3.37) we turn the attention away from interest rate model and investigate a stock model. The payoff function being a put option, i.e.

$$\max(K - S(t), 0)$$

with an exponential Lévy process, that is

$$S(t) = xe^{\sigma L_t}, \quad x = S(0).$$

A reason why we choose the put option, instead of a call option is because we want to make sure that the payoff function is bounded ⁶, because the exponential Lévy process or moments of the Lévy process can explode and tend towards infinity. Furthermore we observe changes in the stock price one time per day, as in the simulation of the endowment policy. The initial value of the stock is set to 6, the volatility $\sigma = 0.5$, guaranteed return equal to 5 and the strike price $K = 8$. Simulating 100,000 times with the α -stable process ($\alpha = 1.9$) yields figure 3.8.

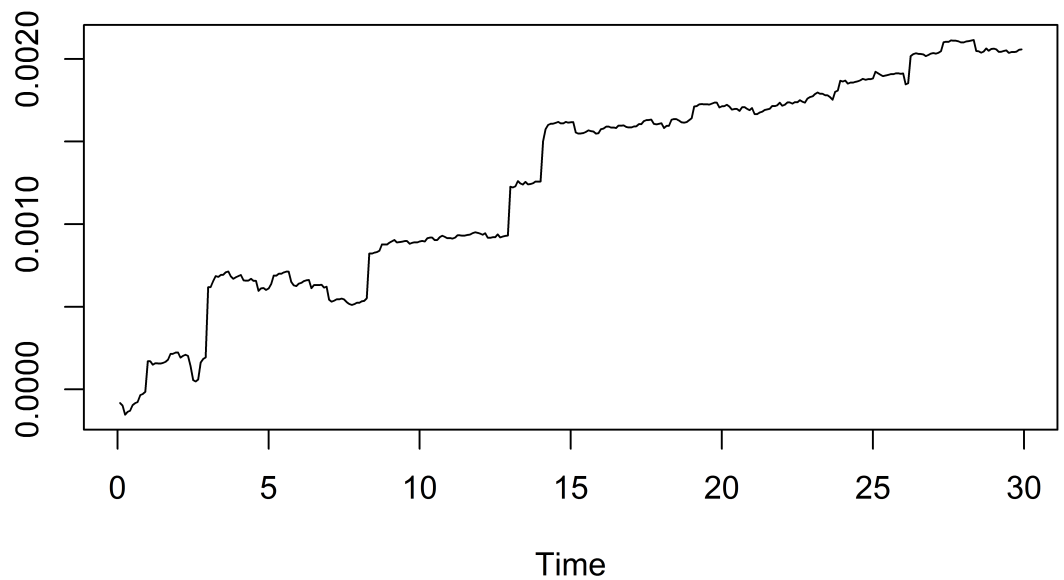


Figure 3.8: Simulation of eq.(3.35) with $\alpha = 1.9$.

We see that the sensitivity of the term life policy with respect to the initial interest rate does not tend to zero, but is rather small. The reason

⁶An alternative option which is bounded is a call option with a roof, so if the stock passes a threshold the top of the profit goes back to the option seller. This option is known as a cliquet option.

why the sensitivity is not tending towards zero is because of the integral. We are actually summing up all the sensitivities at all time points. What about the continuous case, where we choose $\alpha = 2$ so we use the Brownian motion? Choosing the same values as in the previous simulation yields figure 3.9. We observe how smooth the curve is compared to the case when we simulate a jump process (figure 3.8). When conducting simulations it actually turned out to be quite inconsistent, in the sense that the simulation outcome (in case of the Brownian motion) depended heavily on the choice of values for the drift and volatility. One similarity with the α -stable process case is the magnitude of the sensitivity.

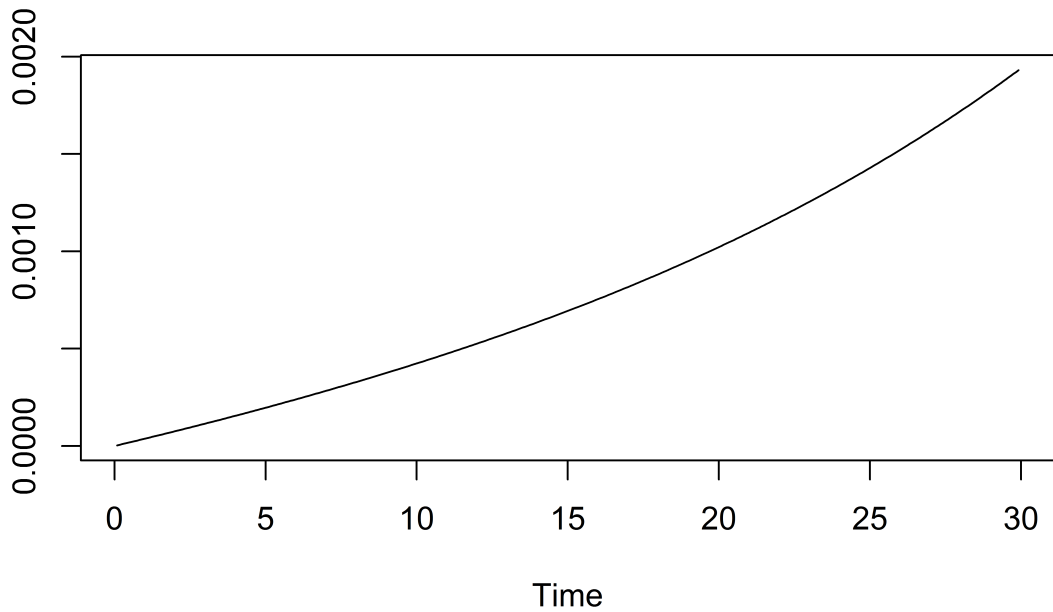


Figure 3.9: Simulation of eq.(3.35) with Brownian motion having drift = 0.1 and volatility = 0.5.

Completing this chapter, we have observed that implementation of Zhang's formula is very interesting and important. We have observed that a heavy

theoretical formula used for computation in sensitivity analysis, yields the results one would expect. We observed how the sensitivity of a caplet, with respect to the initial interest rate tended towards zero. This also became evident as we ran more simulations in the section of application to life insurance using Zhang's formula.

We do however note some limitations in Zhang's formula. The coefficients b in the dynamics (3.9) is assumed to have continuous first order derivative. A class with worse ⁷ coefficients we will investigate in the next chapter are Hölder coefficients. Here we show a Bismut-Elworthy-Li type formula, in the case of the Greek Delta.

⁷In the sense that we do not have as nice properties as a coefficient $b \in C^1(\mathbb{R})$, which is bounded.

Chapter 4

Derivation of the delta in the case of SDE's with Hölder drift coefficients

In this chapter we will derive a "derivative free" representation for the Greek delta (i.e. we will derive a Bismut-Elworthy-Li type formula). This is a new formula which is of the form (3.13), but has the new feature that we will be able to deal with a larger class of coefficients b . The derivation of the formula uses results from [8], as they derive results concerning SDE's on the same form we will work with, that is, with Hölder coefficients. We will use this paper in connection with [17].

The following chapter is divided into three sections, where the first section introduces the framework and states "computational rules" needed further on. Also we include a short summary of the paper [8]. In section two we state two lemmas, with proofs, which are needed for the main theorem of this chapter. The first lemma ensures uniform boundedness of derivatives of a approximating stochastic flows in the L^2 we will work with. Then, in the second lemma, we show that the SDE-solution with Hölder drift coefficient in the forthcoming Bismut-Elworthy-Li formula, is Sobolev differentiable. The important Bismut-Elworthy-Li formula for the Greek Delta with Hölder coefficients is given in a theorem, presented in the last section of this chapter.

4.1 Framework

The main paper used, to be able to derive the forthcoming lemmas and final theorem is the paper [8], where they use a new method to construct unique strong solutions of SDE's with singular coefficients. Let's consider an SDE of the form

$$X_t = x + \int_0^t b(s, X_s) ds + L_t, \quad 0 \leq t \leq T, x \in \mathbb{R}^d \quad (4.1)$$

where $b : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}^d$ is a Borel-measurable function and $L_t, 0 \leq t \leq T$ is a d-dimensional (square integrable) Lévy process. So the process is on some complete probability space $(\Omega, \mathcal{F}, \mu)$, that has stationary and independent increments starting at zero. In fact, we have that b is singular in the sense that b is bounded and β - Hölder continuous, i.e.

$$\|b\|_{C_b^\beta} := \sup_{0 \leq t \leq T, x \in \mathbb{R}^d} |b(t, x)| + \sup_{0 \leq t \leq T} \sup_{x \neq y} \frac{|b(t, x) - b(t, y)|}{|x - y|^\beta} < \infty,$$

where $0 < \beta < 1$. We know (by Picard iteration) that if b is Lipschitz continuous that there exists a unique strong solution $X_t, 0 \leq t \leq T$ to (4.1). This type of solution is used in a variety of applications, such as in statistical mechanics or in the theory of controlled diffusion processes. In the literature there are results when b is singular and L_t is a Wiener process. On the other hand, we have the case when L_t is a pure jump Lévy process. The aim of [8] is to introduce a new technique to construct unique strong solutions to (4.1). The way this is done, is by approximating the singular coefficients b in (4.1) by smooth functions b_n . So

$$X_t^n = x + \int_0^t b_n(s, X_s^n) ds + L_t, \quad 0 \leq t \leq T, \quad x \in \mathbb{R}^d$$

for each $n \geq 1$. Then the authors apply a compactness criterion based on Malliavin calculus in connection with PDE-techniques to the sequence of solutions $X_t^n, n \geq 1$ to obtain a unique strong solution X_t . One important feature of this technique is that the solution X_t will be Malliavin differentiable for all t .

Framework

We now proceed to introduce the framework we are working in, for the remainder of this chapter.

Martingale is in the ordinary sense, i.e.:

Definition 4.1. A filtration (on (Ω, \mathcal{F})) is a family $\mathcal{M} = \{M_t\}_{t \geq 0}$ of σ -algebras $\mathcal{M}_t \subset \mathcal{F}$ such that

$$0 \leq s < t \Rightarrow \mathcal{M}_s \subset \mathcal{M}_t.$$

An n -dimensional stochastic process $\{M_t\}_t$ on (Ω, \mathcal{F}, P) is called a martingale with respect to a filtration $\mathcal{M}_{t \geq 0}$ if

- i) M_t is \mathcal{M}_t -measurable for all t ,
- ii) $E[|M_t|] < \infty$ for all t
- iii) $E[M_s | \mathcal{M}_t] = M_t$ for all $s \geq t$.

We need the Itô isometry in the case of Lévy processes:
Let $X(t) \in \mathbb{R}^n$ be an Itô - Lévy process of the form

$$dX(t) = \alpha(t, \omega)dt + \sigma(t, \omega)dB(t) + \int_{\mathbb{R}^l} \gamma(t, z, \omega)\bar{N}(dt, dz). \quad (4.2)$$

where $\alpha : [0, T] \times \Omega \rightarrow \mathbb{R}^n$, $\sigma : [0, T] \times \Omega \rightarrow \mathbb{R}^{n \times m}$, and $\gamma : [0, T] \times \mathbb{R}^l \times \Omega \rightarrow \mathbb{R}^d$ are adapted processes such that the integrals exist. Here $B(t)$ is an m -dimensional Brownian motion. Also,

$$\bar{N}(dt, dz) = \begin{cases} N(dt, dz) - \nu(dz)dt & \text{if } |z| < R \\ N(dt, dz) & \text{if } |z| \geq R \end{cases}$$

for some $R \in [0, \infty]$.

Theorem 4.2 (The Itô-Lévy isometry). *Let $X(t) \in \mathbb{R}^d$ be as in (4.2) but with $X(0) = 0$ and $\alpha = 0$. Then*

$$\begin{aligned} E[X^2(T)] &= E \left[\int_0^T \left(\sum_{i=1}^n \sum_{j=1}^m \sigma_{ij}^2(t) + \sum_{i=1}^n \sum_{j=1}^l \int_{\mathbb{R}} \gamma_{ij}^2(t, z_j) \nu_j(dz_j) \right) dt \right] \\ &= \sum_{i=1}^n E \left[\int_0^T \left(\sum_{j=1}^m \sigma_{ij}^2(t) + \sum_{j=1}^l \int_{\mathbb{R}} \gamma_{ij}^2(t, z_j) \nu_j(dz_j) \right) dt \right] \end{aligned} \quad (4.3)$$

provided that the right-hand side is finite.

Proof. Theorem 1.17 in [19]. □

Remark. As a special case of Theorem 4.2 assume that

$$dX(t) = d\eta(t) = \int_{\mathbb{R}} z \tilde{N}(dt, dz) \in \mathbb{R}$$

with $E[X^2(T)] = T \int_{\mathbb{R}} z^2 \nu(dz) < \infty$. Then we have the isometry

$$E \left[\left(\int_0^T H(t) d\eta(t) \right)^2 \right] = E \left[\int_0^T H^2(t) dt \right] \int_{\mathbb{R}} z^2 \nu(dz) \quad (4.4)$$

for all $H \in L_{ucp}$ such that $H \in L^2([0, T \times \Omega])$, i.e., such that

$$\|H\|_{L^2([0, T] \times \Omega)} := E \left[\int_0^T H^2(t) dt \right] < \infty.$$

L_{ucp} denotes the space of adapted càglàd processes (left continuous with right limits), for more on this consult [19].

Grönwall inequality

We will use an inequality called the Grönwall inequality ¹.

This inequality states that if we have a non-negative function $v(t)$ such that

$$v(t) \leq C + A \int_0^t v(s) ds, \quad \text{for } 0 \leq t \leq T$$

for constants C, A where $A \geq 0$ then

$$v(t) \leq C \cdot \exp(At), \quad \text{for } 0 \leq t \leq T.$$

Furthermore we will encounter the Sobolev space $W_{loc}^{1,2}$ which consists of the subset of functions $f \in L^2(\mathbb{R})$, such that the derivative up to order 1 exists in the weak sense and belongs to L^2 . More precisely we have the following definition:

Definition 4.3. For a $u \in W_{loc}^{1,2}(\mathbb{R}^d) \Leftrightarrow \forall y \in \mathbb{R}^d \quad \exists U$ open, bounded subset with $y \in U$ such that $u \in W^{1,2}(U)$. Where the space $W^{1,2}$ is defined as

$$W^{1,2} \stackrel{\text{def}}{=} \left\{ u \in L^2(u) : \int_U u(x) \frac{\partial}{\partial x_i} \phi(x) dx = - \int_U f(x) \phi(x) dx \right\}$$

¹If one is interested in how to show that the inequality holds see exercise 5.17 in [18].

for all $\phi \in C_c^\infty(U)$ for some $f \in L^2(U)\}$,

where C_c^∞ denotes the space of functions that has compact support and all orders of the derivatives are continuous.

In the proof we will use the Itô formula ², that is presented in Ikeda and Watanabe, namely Theorem 5.1 in [9]. This is due to the form of the SDE-solution we will work with, where the stochastic part of the SDE-solution is of the form

$$L_t = \int_0^{t^+} \int_{\mathbb{R}^d} z \cdot \mathbf{1}_{\{\|z\| > 1\}} N(ds, dz) \\ + \int_0^{t^+} \int_{\mathbb{R}^d} z \cdot \mathbf{1}_{\{\|z\| \leq 1\}} \tilde{N}(ds, dz).$$

Using the Itô formula presented in chapter 3 in this setting is rather cumbersome, compared to the Itô formula presented in Ikeda and Watanabe.

Furthermore, for $\beta \in (0, 1)$ and $k, d \geq 1$ the space $C_b^\beta(\mathbb{R}^d, \mathbb{R})$ will denote a space of continuous functions $u : \mathbb{R}^d \rightarrow \mathbb{R}$, which we will use on the simplified form $C_b^\beta(\mathbb{R}^d) = C_b^\beta(\mathbb{R}^d, \mathbb{R})$. This space is also known as the space of bounded β - Hölder continuous functions, which is equipped with the norm

$$\|u\|_{C_b^\beta(\mathbb{R}^d)} = \|u\|_\infty + \sup_{x \neq y} \frac{|u(x) - u(y)|}{|x - y|^\beta} < \infty,$$

where $\|u\|_\infty := \sup_{x \in \mathbb{R}^d} |u(x)|$.

Also, $C_\infty(\mathbb{R}^d)$ denotes the space of bounded continuous functions vanishing at infinity, with respect to the supremum norm $\|\cdot\|_\infty$. So $C_\infty^2(\mathbb{R}^d)$ denotes the space of all $f \in C_b^2(\mathbb{R}^d) \cap C_\infty(\mathbb{R}^d)$, so that its first and second order partial derivative belong to $C_\infty(\mathbb{R}^d)$.

In the following sections we will encounter weak convergence ³:

Definition 4.4. *Weak convergence in $L^2(\Omega)$:*

$$X_n \rightharpoonup X \quad \text{in } L^2(\Omega)$$

\Leftrightarrow

$$E[Y X_n] \xrightarrow{n \rightarrow \infty} E[Y \cdot X] \quad \text{for all } Y \in L^2(\Omega).$$

²Note that we work with the Itô-Lévy formula, not the classic Itô formula with the Brownian motion.

³We will later on work with a sequence where we let $X_n \stackrel{\text{def}}{=} \frac{\partial}{\partial x} X_t^{x,n}$.

The proof of Lemma 4.6 is based on Kolmogorov's equation, thus we need the following theorem:

Theorem 4.5. *Let L_t , $0 \leq t \leq T$ be a d -dimensional α -stable process for $\alpha \in (1, 2)$. Require that $\phi \in C([0, T], C_b^\beta(\mathbb{R}^d))$ for $\beta \in (0, 1)$ with $\alpha + \beta > 2$. Then there exist a $u \in C([0, T], C_b^2(\mathbb{R}^d)) \cap C^1([0, T], C_b(\mathbb{R}^d))$ satisfying the backward Kolmogorov equation*

$$\frac{\partial}{\partial t}u + b \cdot \nabla u + \mathcal{L}u = -\phi \quad \text{on } [0, T], \quad u|_{t=T} = 0. \quad (4.5)$$

Moreover

$$\|Du\|_{C_b^\beta} \leq C(T)\|\phi\|_{C_b^\beta} \quad (4.6)$$

where

$$C(T) \rightarrow 0 \quad \text{for } T \searrow 0,$$

as well as

$$\|D^2u\|_\infty \leq M \cdot \|\phi\|_{C_b^\beta} \quad (4.7)$$

for a constant M .

Proof. Theorem 17 in [8]. □

Remark.

$$(\mathcal{L}f)(x) = \int_{\mathbb{R}^d} \{f(x+y) - f(x) - y \cdot \mathbf{1}_{\{\|y\| \leq 1\}} \cdot Df(x)\} \nu(dy),$$

where ν denotes the Lévy measure and $f \in C_\infty^2(\mathbb{R}^d)$.

4.2 Properties of an SDE with Hölder coefficients

In this section we state and show two Lemmas, needed to be able to prove the theorem in the next section.

Lemma 4.6. *Let*

$$X_t^{x,n} = x + \int_0^t b_n(X_s^{x,n}) ds + L_t,$$

where $L_t, 0 \leq t \leq T$ is a d -dimensional α -stable process, where $\alpha \in (1, 2)$ and $b_n \in C([0, T]; C_c^\infty(\mathbb{R}^d))$, where $\beta \in (0, 1)$ such that $\|b_n\|_{C_b^\beta} \leq \|b\|_{C_b^\beta}$ for all n . Furthermore require that $\alpha + \beta > 2$ and $\alpha > 2\beta$ then

$$\sup_{x \in \mathbb{R}^d, n \geq 1} E \left[\left\| \frac{\partial}{\partial x} X_t^{x,n} \right\|^2 \right] < \infty.$$

Proof. Consider now the SDE-solutions

$$X_t^{x,n} = x + \int_0^t b_n(X_s^{x,n}) ds + L_t, \quad 0 \leq t \leq T, \quad n \geq 0, \quad (4.8)$$

where

$$\begin{aligned} L_t &= \int_0^{t^+} \int_{\mathbb{R}^d} z \cdot \mathbf{1}_{\{\|z\| > 1\}} N(ds, dz) \\ &\quad + \int_0^{t^+} \int_{\mathbb{R}^d} z \cdot \mathbf{1}_{\{\|z\| \leq 1\}} \tilde{N}(ds, dz) \end{aligned} \quad (4.9)$$

is an α -stable process such that $\alpha + \beta > 2$ for $\alpha \in (1, 2)$ and $\beta \in (0, 1)$ and where $b_n \in C([0, T]; C_c^\infty(\mathbb{R}^d)), n \geq 1$ such that

$$\|b_n\|_{C_b^\beta} \leq \|b\|_{C_b^\beta}$$

for all n and

$$b_{n_k}(\cdot) \longrightarrow b(\cdot) \quad \text{in } C^\delta(K)$$

for all t , any compact set $K \subset \mathbb{R}^d$ and $0 < \delta < \beta$ for a subsequence $n_k, k \geq 1$ depending on K . Using Theorem 4.5 in connection with Itô's lemma

⁴ applied to $X_t^{x,n}$, $0 \leq t \leq T$, $n \geq 1$ we get ⁵

$$\begin{aligned}
u(t, X_t^{x,n}) &= u(0, x) + \int_0^t \frac{\partial}{\partial s} u(s, X_s^{x,n}) ds \\
&+ \int_0^t \frac{\partial}{\partial x} u(s, X_s^{x,n}) \cdot b(X_s^{x,n}) ds \\
&+ \int_0^{t^+} \int_{\mathbb{R}^d} \{u(s, X_{s^-}^{x,n} + \gamma_1(z)) - u(s, X_{s^-}^{x,n})\} \tilde{N}(ds, dz) \\
&+ \int_0^{t^+} \int_{\mathbb{R}^d} \{u(s, X_{s^-}^{x,n} + \gamma_2(z)) - u(s, X_{s^-}^{x,n})\} \tilde{N}(ds, dz) \\
&+ \int_0^t \int_{\mathbb{R}^d} \{u(s, X_s^{x,n} + \gamma_2(z)) - u(s, X_s^{x,n}) - \gamma_2(z) \cdot D_x u(s, X_s^{x,n})\} \nu(dz) ds \\
&= u(0, x) + \int_0^t b(X_s^{x,n}) ds \\
&+ \int_0^{t^+} \int_{\mathbb{R}^d} \{u(s, X_{s^-}^{x,n} + \gamma_1(z)) - u(s, X_{s^-}^{x,n})\} \tilde{N}(ds, dz) \\
&+ \int_0^{t^+} \int_{\mathbb{R}^d} \{u(s, X_{s^-}^{x,n} + \gamma_2(z)) - u(s, X_{s^-}^{x,n})\} \tilde{N}(ds, dz),
\end{aligned}$$

where

$$\gamma_1(z) := \mathbb{1}_{\{\|z\| > 1\}} \cdot z \quad \text{and} \quad \gamma_2(z) := \mathbb{1}_{\{\|z\| \leq 1\}} \cdot z$$

on \mathbb{R}^d .

So we obtain that

$$\begin{aligned}
X_t^{x,n} &= x + u(t, X_t^{x,n}) - u(0, x) \\
&- \int_0^{t^+} \int_{\mathbb{R}^d} \{u(s, X_{s^-}^{x,n} + \gamma_1(z)) - u(s, X_{s^-}^{x,n})\} \tilde{N}(ds, dz) \\
&- \int_0^{t^+} \int_{\mathbb{R}^d} \{u(s, X_{s^-}^{x,n} + \gamma_2(z)) - u(s, X_{s^-}^{x,n})\} \tilde{N}(ds, dz) \\
&+ L_t
\end{aligned} \tag{4.10}$$

We know from [14] that

$$x \longmapsto X_t^{x,n}$$

⁴The version of Itô's lemma used is the one on page 66 in [9].

⁵Where the function u is the solution to the backward Kolmogorov equation in theorem 4.5

is continuously differentiable for all t, n with probability 1. So taking the partial derivative with respect to x and using the mean value theorem ⁶ we get that

$$\begin{aligned} \frac{\partial}{\partial x} X_t^{x,n} &= Id + Du(t, X_t^{x,n}) \frac{\partial}{\partial x} X_t^{x,n} - Du(0, x) \\ &\quad - \int_0^{t^+} \int_{\mathbb{R}^d} (Du(s, X_{s^-}^{x,n} + \gamma_1(z)) - Du(s, X_{s^-}^{x,n})) \frac{\partial}{\partial x} X_{s^-}^{x,n} \tilde{N}(ds, dz) \\ &\quad - \int_0^{t^+} \int_{\mathbb{R}^d} \left(\int_0^1 D^2 u(s, X_{s^-}^{x,n} + \theta \gamma_2(z)) \frac{\partial}{\partial x} X_{s^-}^{x,n} d\theta \cdot \gamma_2(z) \right) \tilde{N}(ds, dz), \end{aligned} \quad (4.11)$$

where Id denotes the identity matrix. Let $\alpha > 2\beta$ in addition. Then it follows from Theorem 4.5 in connection with the Itô-Lévy isometry that

$$\begin{aligned} E \left[\left\| \frac{\partial}{\partial x} X_t^{x,n} \right\|^2 \right] &\leq K_1 + E \left[\left\| Du(t, X_t^{x,n}) \frac{\partial}{\partial x} X_t^{x,n} \right\|^2 \right] \\ &\quad + E \left[\left\| \int_0^{t^+} \int_{\mathbb{R}^d} \left((Du(s, X_{s^-}^{x,n} + \gamma_1(z)) - Du(s, X_{s^-}^{x,n})) \frac{\partial}{\partial x} X_{s^-}^{x,n} \right) \tilde{N}(ds, dz) \right\|^2 \right] \\ &\quad + E \left[\left\| \int_0^{t^+} \int_{\mathbb{R}^d} \left(\int_0^1 D^2 u(s, X_{s^-}^{x,n} + \theta \gamma_2(z)) \frac{\partial}{\partial x} X_{s^-}^{x,n} d\theta \gamma_2(z) \right) \tilde{N}(ds, dz) \right\|^2 \right] \\ &\leq K_1 + (C(T))^2 \|b_n\|_{C_b^\beta}^2 E \left[\left\| \frac{\partial}{\partial x} X_t^{x,n} \right\|^2 \right] \\ &\quad + \int_0^t E \left[\left\| \frac{\partial}{\partial x} X_s^{x,n} \right\|^2 \right] ds \int_{\mathbb{R}^d} \|Du(s, X_{s^-}^{x,n} + \gamma_1(z)) - Du(s, X_{s^-}^{x,n})\|^2 \nu(dz) \\ &\quad + \int_0^t E \left[\left\| \frac{\partial}{\partial x} X_s^{x,n} \right\|^2 \right] ds \int_{\mathbb{R}^d} \int_0^1 \|D^2 u(s, X_{s^-}^{x,n} + \theta \gamma_2(z))\|^2 d\theta \|\gamma_2(z)\|^2 \nu(dz) \end{aligned}$$

⁶We will use an integral form of the mean value theorem (this form of the mean value theorem is known as Hadamard's mean value theorem), which is of the form $f(x+h) - f(x) = \int_0^1 f'(x + \theta \cdot h) d\theta \cdot h$. Where we let $h = \gamma_1(z)$ in the proof.

$$\begin{aligned}
&\leq K_1 + (C(T))^2 \|b_n\|_{C_b^\beta}^2 E \left[\left\| \frac{\partial}{\partial x} X_t^{x,n} \right\|^2 \right] \\
&+ \int_0^t E \left[\left\| \frac{\partial}{\partial x} X_s^{x,n} \right\|^2 \right] ds \int_{\mathbb{R}^d} \frac{\|Du(s, X_{s-}^{x,n} + \gamma_1(z)) - Du(s, X_{s-}^{x,n})\|^2}{\|\gamma_1(z)\|^{2\beta}} \\
&\cdot \|\gamma_1(z)\|^{2\beta} \nu(dz) \\
&+ \int_0^t E \left[\left\| \frac{\partial}{\partial x} X_s^{x,n} \right\|^2 \right] ds \int_{\mathbb{R}^d} \int_0^1 M^2 \|b_n\|_{C_b^\beta}^2 d\theta \|\gamma_2(z)\|^2 \nu(dz)
\end{aligned}$$

$$\begin{aligned}
&\leq K_1 \{ 1 + (C(T))^2 \cdot \|b_n\|_{C_b^\beta}^2 \cdot E \left[\left\| \frac{\partial}{\partial x} X_t^{x,n} \right\|^2 \right] \right. \\
&+ (C(T))^2 \|b_n\|_{C_b^\beta}^2 \int_{\mathbb{R}^d} \|\gamma_1(z)\|^{2\beta} \nu(dz) \int_0^t E \left[\left\| \frac{\partial}{\partial x} X_s^{x,n} \right\|^2 \right] ds \\
&\left. + M^2 \|b_n\|_{C_b^\beta}^2 \cdot \int_{\mathbb{R}^d} \|\gamma_2(z)\|^2 \nu(dz) \int_0^t E \left[\left\| \frac{\partial}{\partial x} X_s^{x,n} \right\|^2 \right] ds \right\},
\end{aligned}$$

where K_1 is a constant, and

$$E \left[\int_0^t \left\| \frac{\partial}{\partial x} X_{s-}^{x,n} \right\|^2 ds \right] = E \left[\int_0^t \left\| \frac{\partial}{\partial x} X_s^{x,n} \right\|^2 ds \right].$$

So we get that

$$\begin{aligned}
E \left[\left\| \frac{\partial}{\partial x} X_t^{x,n} \right\|^2 \right] &\leq \frac{1}{1 - K_1 (C(T))^2 \|b_n\|_{C_b^\beta}^2} \\
&\cdot K_1 (1 + \tilde{C} \int_0^t E \left[\left\| \frac{\partial}{\partial x} X_s^{x,n} \right\|^2 \right] ds)
\end{aligned}$$

Thus Grönwall's Lemma gives

$$\begin{aligned}
E \left[\left\| \frac{\partial}{\partial x} X_t^{x,n} \right\|^2 \right] &\leq \frac{K_1}{1 - K_1 (C(T))^2 \|b_n\|_{C_b^\beta}^2} \\
&\cdot \exp \left(T \cdot \frac{\tilde{C} \cdot K_1}{1 - K_1 (C(T))^2 \|b_n\|_{C_b^\beta}^2} \right),
\end{aligned}$$

where we choose T such that

$$1 - K_1 (C(T))^2 \|b_n\|_{C_b^\beta}^2 \geq \frac{1}{2}$$

uniformly in n . Hence we obtain that

$$\sup_{x \in \mathbb{R}^d, n \geq 1} E[|\frac{\partial}{\partial x} X_t^{x,n}|^2] < \infty.$$

□

In the sequel we aim at using Lemma 4.6 to prove that X_t^x satisfying

$$X_t^x = x + \int_0^t b(X_s^x) ds + L_t \quad 0 \leq s \leq T$$

is Sobolev differentiable in $x \in \mathbb{R}^d$ for all t and almost all ω , if b is Hölder continuous.

Lemma 4.7. *Assume that $\alpha + \beta > 2$, $\alpha > 2\beta$ for some $\alpha \in (1, 2)$ and $\beta \in (0, 1)$. Furthermore, suppose that $b \in C_b^\beta(\mathbb{R}^d)$. Then the solution X_t^x , $0 \leq t \leq T$ associated with b is Sobolev differentiable, that is*

$$(x \longrightarrow X_t^x(\omega)) \in W_{loc}^{1,2}(\mathbb{R}^d)$$

for all t and $\omega \in \tilde{\Omega}$ with $P(\tilde{\Omega}) = 1$

Proof. Consider the approximating sequence b_n of coefficients to b in Lemma 4.6. It is known in this case that

$$X_{\cdot}^{x,n} \xrightarrow{n \rightarrow \infty} X_{\cdot}^x \quad \text{in} \quad L^2([0, T] \times \Omega) \quad (4.12)$$

(at least for a subsequence) see [16]. Choose now $\phi \in C_c^\infty(U)$ for a bounded open set $U \subseteq \mathbb{R}^d$ and $\tau \in L^\infty([0, T] \times \Omega)$. Then

$$\begin{aligned} E \left[\int_0^T \int_U \frac{\partial}{\partial x_i} X_s^{x,n} \phi(x) dx \quad \tau(s, \omega) ds \right] &\stackrel{(*)}{=} -E \left[\int_0^T \int_U X_s^{x,n} \frac{\partial}{\partial x_i} \phi(x) dx \quad \tau(s, \omega) ds \right] \\ &\stackrel{(4.12)}{\xrightarrow{n \rightarrow \infty}} -E \left[\int_0^T \int_U X_s^x \frac{\partial}{\partial x_i} \phi(x) dx \quad \tau(s, \omega) ds \right], \end{aligned}$$

where the $(*)$ equality follows from integration by parts.

On the other hand we know because of Lemma 4.6 that there exists a subsequence (say n for simplicity) such that

$$((t, x, \omega) \longrightarrow \frac{\partial}{\partial x_i} X_t^{x,n}(\omega)) \quad (4.13)$$

converges weakly in $L^2([0, T] \times u \times \Omega)$ to a process $(t, x, \omega) \longrightarrow Y_t^{i,x}$. So we see that

$$E \left[\int_0^T \int_u \frac{\partial}{\partial x_i} X_s^{x,n} \phi(x) dx \quad \tau(s, \omega) ds \right] \xrightarrow[n \rightarrow \infty]{(4.13)} E \left[\int_0^T \int_u Y_s^{i,x} \phi(x) dx \quad \tau(s, \omega) ds \right].$$

Since ϕ and τ are arbitrary, it follows that

$$Y_t^{i,x} = \frac{\partial}{\partial x_i} X_t^x$$

in the distributional sense, so the proof follows. \square

4.3 A Bismut-Elworthy-Li formula ("delta") with Hölder coefficients

Next we want to show the main theorem, providing us with a Bismut-Elworthy-Li's type formula for the Greek delta in the case of Hölder coefficients. We know from Lemma 4.6 that our process $\frac{\partial}{\partial x} X_t^{x,n}$ is uniformly bounded in $x \in \mathbb{R}^d$ and $n \geq 1$ w.r.t. the L^2 -norm. We also know that X_t^x is Sobolev differentiable in $x \in \mathbb{R}^d$ for all t for $\omega \in \tilde{\Omega}$ by Lemma 4.7. Thus we have established two necessary results to be able to prove the main theorem and the main result of this chapter:

Theorem 4.8. *Let*

$$X_t^x = x + \int_0^t b(X_s^x) ds + L_t,$$

where $L_t, 0 \leq t \leq T$ is a d -dimensional α -stable process and $S_T^{\alpha/2}$ is the $\frac{\alpha}{2}$ -stable subordinator, where $\alpha \in (1, 2)$ and $b \in C_b^\beta(\mathbb{R}^d)$, where $\beta \in (0, 1)$. Furthermore require that $\alpha + \beta > 2$ and $\alpha > 2\beta$, also let $g \in C_b(\mathbb{R}^d)$ then

$$\frac{\partial}{\partial x} E[g(X_T^x)] = E \left[g(X_T^x) \cdot \frac{1}{S_T^{\alpha/2}} \int_0^T \frac{\partial}{\partial x} X_s^x dL_s \right] a.e. \quad (4.14)$$

Proof. We approximate X_T^x by $X_T^{x,n}$ from Lemma 4.6 and show that this converges to (4.14), we will use a standard technique used to show that an object is Sobolev differentiable, namely multiply by a smooth function with compact support and then integrate. Without loss of generality let us consider the case $d = 1$. First off we know from Zhang [17] that

$$\frac{\partial}{\partial x} E[g(X_T^{x,n})] = E[g(X_T^{x,n}) \frac{1}{S_T^{\alpha/2}} \int_0^T \frac{\partial}{\partial x} X_s^{x,n} dL_s]. \quad (4.15)$$

Let ⁷ $\delta \in C_c^\infty(U)$, where U is a bounded open subset of \mathbb{R} . Then

$$\begin{aligned} \int_U \frac{\partial}{\partial x} E[g(X_T^{x,n})] \delta(x) dx &\stackrel{(*)}{=} - \int_U E[g(X_T^{x,n})] \frac{\partial}{\partial x} \delta(x) dx \\ &\xrightarrow[\text{convergence}]{\text{dominated}} - \int_U E[g(X_T^x)] \frac{\partial}{\partial x} \delta(x) dx \quad \text{for } n \rightarrow \infty. \end{aligned} \quad (4.16)$$

Where the $(*)$ equality follows by using the classical integration by parts formula, where $E[g(X_T^{x,n})] \delta(x)$ vanishes at the limits i.e. this expression becomes zero, because of the fact that $\delta(x)$ is a smooth function with compact support.

On the other hand,

$$\begin{aligned} E[g(X_T^{x,n}) \frac{1}{S_T^{\alpha/2}} \int_0^T \frac{\partial}{\partial x} X_s^{x,n} dL_s] = & \underbrace{E[(g(X_T^{x,n}) - g(X_T^x)) \frac{1}{S_T^{\alpha/2}} \int_0^T \frac{\partial}{\partial x} X_s^{x,n} dL_s]}_{I_1} \\ & + \underbrace{E[g(X_T^x) \frac{1}{S_T^{\alpha/2}} \int_0^T \frac{\partial}{\partial x} X_s^{x,n} dL_s]}_{I_2} \end{aligned}$$

We see that

$$\begin{aligned} I_1 = & \underbrace{E[(g(X_T^{x,n}) - g(X_T^x)) \frac{1}{S_T^{\alpha/2}} \int_0^{T^+} \int_{\mathbb{R}} \frac{\partial}{\partial x} X_s^{x,n} \cdot z \cdot \mathbf{1}_{\{|z|>1\}} N(ds, dz)]}_{I_{11}} \\ & + \underbrace{E[(g(X_T^{x,n}) - g(X_T^x)) \frac{1}{S_T^{\alpha/2}} \int_0^{T^+} \int_{\mathbb{R}} \frac{\partial}{\partial x} X_s^{x,n} \cdot z \cdot \mathbf{1}_{\{|z|\leq 1\}} \tilde{N}(ds, dz)]}_{I_{12}} \end{aligned}$$

Case I_{11} : choose $p, q > 1$ such that $q < \alpha$ and

$$\frac{1}{p} + \frac{1}{q} = 1$$

⁷The function δ is called a test function.

Then Hölder's inequality ⁸ gives

$$I_{11} \leq \underbrace{E \left[|g(X_T^{x,n}) - g(X_T^x)|^p \cdot \frac{1}{|S_T^{\alpha/2}|^p} \right]}_{I_{111}}^{\frac{1}{p}} \cdot \underbrace{E \left[\left| \int_0^{T^+} \int_{\mathbb{R}} \frac{\partial}{\partial x} X_s^{x,n} \cdot z \cdot \mathbb{1}_{\{\|z\|>1\}} N(ds, dz) \right|^q \right]}_{I_{112}}^{\frac{1}{q}}$$

Case I_{111} : Using dominated convergence and the fact that all moments of $\frac{1}{S_T^{\alpha/2}}$ exist (see Zhang [17]), we obtain that

$$I_{111} \xrightarrow{n \rightarrow \infty} 0$$

Case I_{112} : It follows from e.g. Lemma 8.22 in [13] that

$$\begin{aligned} I_{112} &\leq C_q \{ E \left[\int_0^T \int_{\mathbb{R}} \left| \frac{\partial}{\partial x} X_s^{x,n} \right|^q \cdot |z|^q \cdot \mathbb{1}_{\{\|z\|>1\}} \nu(dz) ds \right]^{\frac{1}{q}} \\ &\quad + E \left[\left| \int_0^T \int_{\mathbb{R}} \frac{\partial}{\partial x} X_s^{x,n} \cdot z \cdot \mathbb{1}_{\|z\|>1} \nu(dz) ds \right|^q \right]^{\frac{1}{q}} \} \\ &= C_q \left\{ \left(\int_{\mathbb{R}} |z|^q \cdot \mathbb{1}_{\{\|z\|>1\}} \nu(dz) \right)^{\frac{1}{q}} \cdot E \left[\int_0^T \left| \frac{\partial}{\partial x} X_s^{x,n} \right|^q ds \right]^{\frac{1}{q}} \right. \\ &\quad \left. + \left| \int_{\mathbb{R}} z \mathbb{1}_{\|z\|>1} \nu(dz) \right| E \left[\left| \int_0^T \frac{\partial}{\partial x} X_s^{x,n} ds \right|^q \right]^{\frac{1}{q}} \right\} \\ &\stackrel{q < \alpha < 2}{\leq} C < \infty, \end{aligned}$$

where C_q is a constant. So this implies that

$$I_{11} \xrightarrow{n \rightarrow \infty} 0$$

⁸For Hölder's inequality see e.g. Theorem 1.55 in [15].

Case I_{12} :

$$\begin{aligned}
I_{12} \leq & \underbrace{E \left[|g(X_T^{x,n}) - g(X_T^x)|^2 \cdot \frac{1}{|S_T^{\alpha/2}|^2} \right]}_{I_{121}}^{\frac{1}{2}} \\
& \cdot \underbrace{E \left[\left| \int_0^{T^+} \int_{\mathbb{R}} \frac{\partial}{\partial x} X_s^{x,n} \cdot z \cdot \mathbf{1}_{\{\|z\| \leq 1\}} \tilde{N}(ds, dz) \right|^2 \right]}_{I_{122}}^{\frac{1}{2}} \quad (4.17)
\end{aligned}$$

Case I_{121} : Dominated convergence yields

$$I_{121} \xrightarrow{n \rightarrow \infty} 0$$

Case I_{122} : Itô's isometry gives

$$\begin{aligned}
I_{122} &= E \left[\int_0^T \int_{\mathbb{R}} \left| \frac{\partial}{\partial x} X_s^{x,n} \right|^2 \cdot |z|^2 \cdot \mathbf{1}_{\{\|z\| \leq 1\}} \nu(dz) ds \right]^{\frac{1}{2}} \\
&= \left(\int_{\mathbb{R}} |z|^2 \cdot \mathbf{1}_{\{\|z\| \leq 1\}} \nu(dz) \right)^{\frac{1}{2}} \cdot E \left[\int_0^T \left| \frac{\partial}{\partial x} X_s^{x,n} \right|^2 ds \right]^{\frac{1}{2}} \\
&\leq C, \quad (4.18)
\end{aligned}$$

where we have used the Itô isometry of the form

$$E \left[\left(\int_0^T \int_{\mathbb{R}_0} f(s, z) \tilde{N}(ds, dz) \right)^2 \right] = E \left[\int_0^T \int_{\mathbb{R}_0} (f(s, z))^2 \nu(dz) ds \right].$$

So we see that

$$I_{12} \xrightarrow{n \rightarrow \infty} 0$$

So it follows that

$$I_1 \xrightarrow{n \rightarrow \infty} 0$$

Case I_2 : We can write I_2 as

$$\begin{aligned}
I_2 &= E \left[\underbrace{g(X_T^x) \frac{1}{S_T^{\alpha/2}} \int_0^{T^+} \int_{\mathbb{R}} \frac{\partial}{\partial x} X_s^{x,n} \cdot z \cdot \mathbf{1}_{\{\|z\| \leq 1\}} \tilde{N}(ds, dz)}_{I_{21}} \right] \\
&+ E \left[\underbrace{g(X_T^x) \cdot \frac{1}{S_T^{\alpha/2}} \int_0^{T^+} \int_{\mathbb{R}} \frac{\partial}{\partial x} X_s^{x,n} \cdot z \cdot \mathbf{1}_{\{\|z\| > 1\}} N(ds, dz)}_{I_{22}} \right]
\end{aligned}$$

Case I_{21} : Set

$$\xi = g(X_T^x) \cdot \frac{1}{S_T^{\alpha/2}} \in L^p(P), \quad \text{for } p > 2.$$

Then Itô's representation theorem implies that

$$\xi = E[\xi] + \int_0^T \int_{\mathbb{R}} \psi(s, z) \tilde{N}(ds, dz)$$

for a unique predictable process $\psi \in L^2(ds \times \nu(dz) \times dP)$

So

$$\begin{aligned} I_{21} &= E \left[\underbrace{E[\xi] \cdot \int_0^T \int_{\mathbb{R}} \frac{\partial}{\partial x} X_s^{x,n} \cdot z \cdot \mathbf{1}_{\{\|z\| \leq 1\}} \tilde{N}(ds, dz)}_{\text{martingale}} \right] \\ &\quad + E \left[\int_0^T \int_{\mathbb{R}} \psi(s, z) \tilde{N}(ds, dz) \cdot \int_0^T \int_{\mathbb{R}} \frac{\partial}{\partial x} X_s^{x,n} \cdot z \cdot \mathbf{1}_{\{\|z\| \leq 1\}} \tilde{N}(ds, dz) \right] \\ &= E \left[\int_0^T \int_{\mathbb{R}} \psi(s, z) \tilde{N}(ds, dz) \cdot \int_0^T \int_{\mathbb{R}} \frac{\partial}{\partial x} X_s^{x,n} \cdot z \cdot \mathbf{1}_{\{\|z\| \leq 1\}} \tilde{N}(ds, dz) \right] \\ &\stackrel{(*)}{=} E \left[\int_0^T \int_{\mathbb{R}} \psi(s, z) \cdot \frac{\partial}{\partial x} X_s^{x,n} \cdot z \cdot \mathbf{1}_{\{\|z\| \leq 1\}} \nu(dz) ds \right] \\ &= E \left[\int_0^T \underbrace{\int_{\mathbb{R}} \psi(s, z) \cdot z \cdot \mathbf{1}_{\{\|z\| \leq 1\}} \nu(dz)}_{\in L^2(ds \times dP)} \frac{\partial}{\partial x} X_s^{x,n} ds \right], \end{aligned}$$

where the $(*)$ equality follows from the Itô isometry ⁹.

Since $\frac{\partial}{\partial x} X_s^{x,n}$ is weakly convergent in $L^2(ds \times dx \times dP)$ (for a subsequence, say n) to $\frac{\partial}{\partial x} X_s^x$ we see that

$$\int_u I_{21} \delta(x) dx \xrightarrow{n \rightarrow \infty} \int_u E \left[\int_0^T \int_{\mathbb{R}} \psi(s, z) \cdot z \cdot \mathbf{1}_{\{\|z\| \leq 1\}} \nu(dz) \frac{\partial}{\partial x} X_s^x ds \right] \delta(x) dx$$

⁹The technique used here is known as polarization, the idea is to look at two functions f and g then work with $E[(\int f - g)^2] - E[(\int f + g)^2]$ and then apply the Itô isometry to show that $\int f \cdot \int g = \int f \cdot g$, where we have avoided all the technical details, to illustrate the idea.

Case I_{22} : Using Itô's representation theorem we get

$$I_{22} = \underbrace{E \left[E[\xi] \cdot \int_0^T \int_{\mathbb{R}} \frac{\partial}{\partial x} X_s^{x,n} \cdot z \cdot \mathbf{1}_{\{|z|>1\}} N(ds, dz) \right]}_{I_{221}} + \underbrace{E \left[\int_0^T \int_{\mathbb{R}} \psi(s, z) \tilde{N}(ds, dz) \int_0^T \int_{\mathbb{R}} \frac{\partial}{\partial x} X_s^{x,n} \cdot z \cdot \mathbf{1}_{\{|z|>1\}} N(ds, dz) \right]}_{I_{222}}$$

Case I_{221} :

$$\begin{aligned} I_{221} &= E[\xi] \cdot E \left[\int_0^T \int_{\mathbb{R}} \frac{\partial}{\partial x} X_s^{x,n} \cdot z \cdot \mathbf{1}_{\{|z|>1\}} \nu(dz) ds \right] \\ &= E[\xi] \cdot \int_{\mathbb{R}} z \cdot \mathbf{1}_{\{|z|>1\}} \nu(dz) \cdot E \left[\int_0^T \frac{\partial}{\partial x} X_s^{x,n} ds \right] \end{aligned}$$

Thus, by weak convergence of $\frac{\partial}{\partial x} X_s^{x,n}$ we obtain that

$$\int_U I_{221} \delta(x) dx \xrightarrow{n \rightarrow \infty} E[\xi] \cdot \int_{\mathbb{R}} z \cdot \mathbf{1}_{\{|z|>1\}} \nu(dz) \int_U E \left[\int_0^T \frac{\partial}{\partial x} X_s^x ds \right] \delta(x) dx$$

Case I_{222} : It follows from Itô's formula (see e.g. page 66 in Ikeda, Watanabe¹⁰ [9]) that

$$\begin{aligned} \int_0^T \int_{\mathbb{R}} \psi(s, z) \tilde{N}(ds, dz) \cdot \int_0^T \int_{\mathbb{R}} \frac{\partial}{\partial x} X_s^{x,n} \cdot z \cdot \mathbf{1}_{\{|z|>1\}} N(ds, dz) &= \\ &= \int_0^{T^+} \int_{\mathbb{R}} \int_0^{s^-} \int_{\mathbb{R}} \psi(u, \tilde{z}) \tilde{N}(du, d\tilde{z}) \cdot \frac{\partial}{\partial x} X_s^{x,n} \cdot z \cdot \mathbf{1}_{\{|z|>1\}} N(ds, dz) \\ &+ \underbrace{\int_0^{T^+} \int_{\mathbb{R}} \int_0^{s^-} \int_{\mathbb{R}} \frac{\partial}{\partial x} X_u^{x,n} \cdot \tilde{z} \cdot \mathbf{1}_{\{|z|>1\}} N(du, d\tilde{z}) \psi(s, z) \tilde{N}(ds, dz)}_{\text{local martingale}} \end{aligned}$$

¹⁰This is straight forward computation using a 2-dimensional process, i.e. we let $F(X_1, X_2) = X_1 \cdot X_2$, where X_1 denotes the first double integral and X_2 denotes the second double integral.

So using stopping time localization applied to local martingales, we get that

$$\begin{aligned}
I_{222} &= E \left[\int_0^{T^+} \int_{\mathbb{R}} \int_0^{s^-} \int_{\mathbb{R}} \psi(u, \tilde{z}) \tilde{N}(du, d\tilde{z}) \cdot \frac{\partial}{\partial x} X_s^{x,n} \cdot z \cdot \mathbf{1}_{\{\|z\|>1\}} N(ds, dz) \right] \\
&= E \left[\int_0^T \int_{\mathbb{R}} \int_0^s \int_{\mathbb{R}} \psi(u, \tilde{z}) \tilde{N}(du, d\tilde{z}) \cdot \frac{\partial}{\partial x} X_s^{x,n} \cdot z \cdot \mathbf{1}_{\{\|z\|>1\}} \nu(dz) ds \right] \\
&= \int_{\mathbb{R}} z \cdot \mathbf{1}_{\{\|z\|>1\}} \nu(dz) \cdot E \left[\underbrace{\int_0^T \int_0^s \int_{\mathbb{R}} \psi(u, \tilde{z}) \tilde{N}(du, d\tilde{z}) \cdot \frac{\partial}{\partial x} X_s^{x,n} ds}_{\in L^2(ds \times dP)} \right].
\end{aligned}$$

Hence, by weak convergence we get

$$\int_U I_{222} \delta(x) dx \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}} z \cdot \mathbf{1}_{\{\|z\|>1\}} \nu(dz) \int_U E \left[\int_0^T \int_0^s \psi(u, \tilde{z}) \tilde{N}(du, d\tilde{z}) \cdot \frac{\partial}{\partial x} X_s^x ds \right] \delta(x) dx.$$

So by multiplying (4.15) by $\delta \in C_c^\infty(U)$ on both sides and by integration on both sides, it follows from the above considerations that

$$\begin{aligned}
- \int_U E[g(X_T^x)] \frac{\partial}{\partial x} \delta(x) dx &= \\
&= \int_U E \left[\int_0^T \int_{\mathbb{R}} \psi(s, z) \cdot z \cdot \mathbf{1}_{\{\|z\| \leq 1\}} \nu(dz) \frac{\partial}{\partial x} X_s^x ds \right] \delta(x) dx \\
&+ E[\xi] \cdot \int_{\mathbb{R}} z \cdot \mathbf{1}_{\{\|z\|>1\}} \nu(dz) \int_U E \left[\int_0^T \frac{\partial}{\partial x} X_s^x ds \right] \delta(x) dx \\
&+ \int_{\mathbb{R}} z \cdot \mathbf{1}_{\{\|z\|>1\}} \nu(dz) \int_U E \left[\int_0^T \int_0^s \psi(u, \tilde{z}) \tilde{N}(du, d\tilde{z}) \cdot \frac{\partial}{\partial x} X_s^x ds \right] \delta(x) dx.
\end{aligned}$$

By using the same arguments in a reversed way we finally get that

$$- \int_U E[g(X_T^x)] \frac{\partial}{\partial x} \delta(x) dx = \int_U E \left[g(X_T^x) \cdot \frac{1}{S_T^{\alpha/2}} \int_0^T \frac{\partial}{\partial x} X_s^x dL_s \right] \delta(x) dx$$

for all $\delta \in C_c^\infty(U)$.

So $(x \rightarrow E[g(X_T^x)]) \in W_{loc}^{1,2}(\mathbb{R})$ (Sobolev differentiable) with

$$\frac{\partial}{\partial x} E[g(X_T^x)] = E \left[g(X_T^x) \cdot \frac{1}{S_T^{\alpha/2}} \int_0^T \frac{\partial}{\partial x} X_s^x dL_s \right]$$

a.e.

□

Chapter 5

Extensions

In chapter 2 we gave an overview of some basic methods used to compute Greeks. In chapter 3 we implemented Zhang's formula in connection with the Vasicek model, here we saw proof that Zhang's formula indeed works, not just as a theoretical formula in a paper. This was due to the fact that we looked at the sensitivity of a caplet, with respect to the initial interest rate, which went to zero as time passed by, as it should. One could extend the simulation by making them more realistic, e.g. for the Vasicek model one could apply a regime-switching mean reversion rate. This is more realistic because the convergence rate back to the mean might be different if the interest rate is below or above a certain threshold. More precisely, replace the coefficient a in (3.15) by

$$a_1 \mathbb{1}_{\{r(t) > R\}} + a_2 \mathbb{1}_{\{r(t) \leq R\}},$$

where $r(t)$ is the interest rate at time t ($0 \leq t \leq T$) and R a given threshold. This extension seems to be challenging, since the drift coefficient becomes discontinuous. The latter extension would make the previous simulations of Zhang's formula more realistic, and even more realistic if they were fitted to data from the real world ¹. The most prominent extension to this thesis would be to implement the application of Zhang's formula with stochastic interest rate (for the discount factor), namely to unit linked policies. Recall that the previous simulation of a caplet had a deterministic discount factor, a more realistic scenario would be to introduce a stochastic interest rate model for this factor.

¹When taking the derivative to e.g. an initial interest rate we would expect to get zero as time increases, hence the previous results would be about the same even if we choose an improved version of the Vasicek model

A drastic change to the previous simulation would be to change the Vasicek interest rate model to a perhaps better interest rate model. One needs to be careful in this process, as one of the assumptions of Zhang's formula is that the SDE we work with demands that the volatility term is constant, that is it cannot depend on the process itself. One of the perhaps tremendous extensions to this thesis would be to investigate if it is possible to obtain a Bismut-Elworthy-Li type formula for an SDE of the form

$$dX_t(x) = b(t, X_t(x))dt + \sigma(X_t(x))dL_t, \quad X_0(x) = x.$$

This would allow for an implementation of a great class of SDE's and thus more models, which would be interesting for applications in life insurance and other relevant areas.

In this thesis there is presented an extension of Zhang's formula to Hölder coefficients, in the case of the Greek delta. An extension of this would be to investigate if it is possible to achieve a similar result in the case of even worse coefficients. That is, to derive a Bismut-Elworthy-Li type formula for a more general class of coefficients than Hölder ones.

Appendix A

Appendix

A.1 Calculations

We need the Itô formula in order to solve the geometric Brownian motion and the Vasciek model with Brownian motion:

Theorem A.1. *(The 1-dimensional Itô formula) Let $X(t)$ be an Itô process given by*

$$dX(t) = udt + vdB_t.$$

Let $g(t, x) \in C^2([0, \infty) \times \mathbb{R})$ (i.e. g is twice continuously differentiable on $[0, \infty) \times \mathbb{R}$). Then

$$Y_t = g(t, X_t)$$

is again an Itô process, and

$$dY_t = \frac{\partial g}{\partial t}(t, X_t)dt + \frac{\partial g}{\partial x}(t, X_t)dX_t + \frac{1}{2} \frac{\partial^2 g}{\partial x^2}(t, X_t) \cdot (dX_t)^2,$$

where $(dX_t)^2 = (dX_t) \cdot (dX_t)$ is computed according to the rules

$$dt \cdot dt = dt \cdot dB_t = 0, \quad dB_t \cdot dB_t = dt.$$

Proof. See proof of theorem 4.1.2 in [18] □

Lemma A.2 (Solution of equation (2.3)). *The solution of the dynamics*

$$dX(t) = \mu X(t)dt + \sigma X(t)dB(t)$$

where $\mu, \sigma \in \mathbb{R}$ is given by

$$X_t = x \cdot \exp \left(\left(r - \frac{1}{2} \sigma^2 \right) t + \sigma B_t \right).$$

Proof.

$$\begin{aligned} dX(t) &= \mu X(t)dt + \sigma X(t)dB(t) \\ \frac{dX(t)}{X(t)} &= \mu dt + \sigma dB(t) \end{aligned}$$

We recognize that $\frac{dX(t)}{X(t)}$ could originate from the derivative of $\ln(X(t))$, so let's try with Itô formula and see where this leads us.

$$\begin{aligned} d(\ln(X(t))) &= 0dt + \frac{dX(t)}{X(t)} - \frac{1}{2} \frac{(dX(t))^2}{(X(t))^2} \\ &= \frac{dX(t)}{X(t)} - \frac{1}{2} \frac{\sigma^2 (X(t))^2 dt}{(X(t))^2} \\ &= \frac{dX(t)}{X(t)} - \frac{1}{2} \sigma^2 dt \end{aligned}$$

Thus

$$\begin{aligned} d(\ln(X(t))) &= (\mu - \frac{1}{2}\sigma^2)dt + \sigma dB_t \\ \ln \frac{X(t)}{X(0)} &= (\mu - \frac{1}{2}\sigma^2)t + \sigma B(t) \\ X(t) &= x \cdot \exp\{(\mu - \frac{1}{2}\sigma^2)t + \sigma B(t)\} \end{aligned}$$

where we have used that $X(0) = x$.

□

Lemma A.3. *The solution of the dynamics*

$$dr_t = a(b - r_t)dt + \sigma dB_t, \quad r(0) = r_0,$$

is given by

$$r_t = r_0 e^{-at} + b(1 - e^{-at}) + \sigma e^{-at} \int_0^t e^{as} dB_s.$$

Proof. With eq.(3.15) where B_t denotes the Brownian motion, so $B_t \sim N(0, t)$, we can use Itô's lemma to solve the Vasicek dynamics, using the integrating factor technique, we let $g(t, X_t) = e^{at} r_t$ in the Itô formula and

get the following:

$$\begin{aligned}
d(e^{at}r_t) &= ae^{at}r_tdt + e^{at}d(r_t) \\
&= ae^{at}r_tdt + e^{at}(a(b - r_t)dt + \sigma dB_t) \\
&= abe^{at} + e^{at}\sigma dB_t \\
e^{at}r_t - e^0r_0 &= ab \int_0^t e^{as}ds + \sigma \int_0^t e^{as}dB_s \\
e^{at}r_t &= r_0 + ab\left(\frac{1}{a}e^{at} - \frac{1}{a}e^0\right) + \sigma \int_0^t e^{as}dB_s \\
r_t &= r_0e^{-at} + b(1 - e^{-at}) + \sigma e^{-at} \int_0^t e^{as}dB_s
\end{aligned}$$

□

Lemma A.4. *Using eq.(2.9) the approximation for the second derivative*

$$\Gamma = \frac{\partial^2 u(x)}{\partial x^2}$$

becomes

$$\Gamma = \frac{\partial^2 u(x)}{\partial x^2} \approx \frac{u(x + \epsilon^*) - 2u(x) + u(x - \epsilon^*)}{\epsilon^{*2}}.$$

Proof.

$$\begin{aligned}
\Gamma &= \frac{\partial^2 u(x)}{\partial x^2} \\
&= \frac{\partial}{\partial x} \left(\frac{u(x + \epsilon) - u(x - \epsilon)}{2\epsilon} \right) \\
&= \frac{1}{2\epsilon} \left(\frac{\partial}{\partial x} u(x + \epsilon) - \frac{\partial}{\partial x} u(x - \epsilon) \right) \\
&= \frac{1}{4\epsilon^2} (u(x + 2\epsilon) - 2u(x) + u(x - 2\epsilon)) \\
&= \frac{u(x + \epsilon^*) - 2u(x) + u(x - \epsilon^*)}{\epsilon^{*2}}
\end{aligned}$$

where $\epsilon^* = 2\epsilon$

□

A.2 R-code for simulations

Listing A.1: Simulation of α -stable process figure 3.1

```
#Simulation of trajectory for symmetric alpha stable
process
setwd("C:/Users/Tor_Martin/Documents")

png("44alpha.png", width=15, height=15, units='cm', res
    =1500)

n = 1000
alpha = c(0.1,0.5,1,1.9)
y = matrix(0,length(alpha),n)
for(j in 1:length(alpha)){

  t = c()
  t[1] = 0
  t[n] = 100
  spacing = (t[n]-t[1])/n

  gamma = runif(n, -pi/2, pi/2)
  W = rexp(n)
  deltaX =c()
  X = c()

  for (i in 1:n){
    t[i+1] = t[i] + spacing
    deltaX[i] = (t[i+1] - t[i])^(1/alpha[j])*sin(alpha[j]*
      gamma[i+1])/(cos(gamma[i+1]))^(1/alpha[j])*
    (cos((1-alpha[j])*gamma[i+1])/W[i+1])^((1-alpha[j])/
      alpha[j])

    X[i] = sum(deltaX[1:i])
  }
  y[j,]= X
}
```



```

par(mfrow=c(2,2))
plot(t[1:n-1],y[1,(1:n-1)], type = "l", ylab = "", xlab
     = "", main = "Alpha_ = 0.1")
plot(t[1:n-1],y[2,(1:n-1)], type = "l", ylab = "", xlab
     = "", main = "Alpha_ = 0.5" )
plot(t[1:n-1],y[3,(1:n-1)], type = "l", ylab = "", xlab
     = "", main = "Alpha_ = 1")
plot(t[1:n-1],y[4,(1:n-1)], type = "l", ylab = "", xlab
     = "", main = "Alpha_ = 1.9")

dev.off()

```

Listing A.2: Simulation of Vasicek figure 3.2

```

#Vasicek model with Brownian motion

n = 500           #Number of points in Vasicek
K = 0.03          #Strike of the caplet/floorlet

r0 = 0.044        # Initial interest rate
a = 0.05          # Will depend on monetary policy
b = 0.04          # Teta: Longterm interest rate
sigma = 0.004     #Volatility

m = 10000         #Number of simulations for expectation
counter = 1
process = matrix(0,m,n)
for(j in 1:m){
  r = c()
  r[1] = r0

#Simulation of trajectory for Brownian motion

```

```

sigma_BM = 0.1  #Volatility of Brownian motion
beta = 0        #Drift of Brownian motion
t = c()
t[1] = 0        #t0 = t[1]
t[n] = 500
spacing = (t[n]-t[1])/(n)

for (i in 2:(n+1)){
t[i] = t[i-1] + spacing
}

N = rnorm(n)
deltaX = c()
L = c()
for(i in 1:n){
deltaX[i] = sigma_BM*N[i]*sqrt(spacing) + beta*spacing
L[i] = sum(deltaX[1:i])
}

#Vasicek with BM

delta_t = spacing

for(i in 1:(n-1)){
dr = a*(b-r[i])*delta_t + sigma*(L[i+1]-L[i])
r[i+1] = r[i] + dr
}

process[j,] = r
print(counter)
counter = counter +1
}

expectation = c()

for(i in 1:n){
expectation[i] = 1/m*sum(process[,i])
}

```

```

}

setwd("C:/Users/Tor_Martin/Documents")

png("vasicek_run.png", width=16, height=10, units='cm',
    res=1500)

plot(t[2:(n+1)], expectation, type = "l", ylim = c
      (0.036,0.045), xlab = "Time", ylab = "r",
main = "Vasicek with a=0.05, b=0.04, r0=0.044 and
      sigma=0.004", lwd = 2)
lines(t[2:(n+1)], process[1,], col = "red")
lines(t[2:(n+1)], process[2,], col = "blue")
lines(t[2:(n+1)], process[3,], col = "green")
lines(t[2:(n+1)], process[4,], col = "purple")

dev.off()

```

Listing A.3: Simulation of Vasicek figure 3.3

```

#Vasicek model with jumps

n = 500 #Number of points in Vasicek

r0 = 0.044 # Initial interest rate
a = 0.05 # Will depend on monetary policy
b = 0.04 # Teta: Longterm interest rate
sigma = 0.004 # Volatility

m = 10000 #Number of simulations for expectation
counter = 1
process = matrix(0,m,n)
for(j in 1:m){
  r = c()
  r[1] = r0

```

```

#Simulation of trajectory for symmetric alpha stable
process
alpha = 1.9

t = c()
t[1] = 0                                #t0 = t[1]
t[n] = 600
spacing = (t[n]-t[1])/(n)

for (i in 2:(n+1)){
t[i] = t[i-1] + spacing
}

gamma = runif(n, -pi/2, pi/2)
W = rexp(n)
deltaX = c()
deltaY = c()
L = c()
St = c()

for (i in 1:n){
deltaX[i] = (spacing)^(1/alpha)*(sin(alpha*gamma[i]))/
((cos(gamma[i]))^(1/alpha))*
(cos((1-alpha)*gamma[i])/W[i])^((1-alpha)/alpha))

deltaY[i] = (spacing)^(1/(0.5*alpha))*(sin((0.5*alpha)*
gamma[i]))/((cos(gamma[i]))^(1/(0.5*alpha)))*
(cos((1-(0.5*alpha))*gamma[i])/W[i])^((1-(0.5*alpha))/
(0.5*alpha)))

L[i] = sum(deltaX[1:i])
St[i] = sum(deltaY[1:i])
}

```

```

#Vasicek with jumps

delta_t = spacing

for(i in 1:(n-1)){
dr = a*(b-r[i])*delta_t + sigma*(L[i+1]-L[i])
r[i+1] = r[i] + dr
}

process[j,] = r
print(counter)
counter = counter +1
}

expectation = c()

for(i in 1:n){
expectation[i] = 1/m*sum(process[,i])
}

setwd("C:/Users/Tor_Martin/Documents")

png("vasicek_run_jump.png", width=16, height=10, units=
'cm', res=1500)

plot(t[2:(n+1)],expectation,type="l", ylim = c
(0.005,0.098), xlab = "Time", ylab = "r",
main = "Vasicek with a=0.05, b=0.04, r0=0.044 and
sigma=0.004",lwd = 2)
lines(t[2:(n+1)],process[7,], col = "red")

dev.off()

```

```

\begin{lstlisting}[frame = single , caption = Simulation
of figure \ref{fig:zhang}]

#Vasicek (Zhang)

n = 600 #Number of points in Vasicek
K = 0.05 #Strike of the caplet (should be higher then
longterm interest rate)

r0 = 0.048      # Initial intrest rate
a = 0.05        # Will depend on monetary policy
b = 0.04        # Teta: Longterm intrest rate
sigma = 0.004   #Volatility

m = 1000000      #Number of simulations for expectation
counter = 1
process = matrix(0,m,n)
for(j in 1:m){
r = c()
r[1] = r0

#Simulation of trajectory for symmetric alpha stable
process
alpha = 0.8

t = c()
t[1] = 0          #t0 = t[1]
t[n] = 600
spacing = (t[n]-t[1])/(n)

for (i in 2:(n+1)){
t[i] = t[i-1] + spacing
}

```

```

gamma = runif(n, -pi/2, pi/2)
W = rexp(n)
deltaX = c()
deltaY = c()
L = c()
St = c()

for (i in 1:n){
deltaX[i] = (spacing)^(1/alpha)*(sin(alpha*gamma[i]))/
((cos(gamma[i]))^(1/alpha))*
(cos((1-alpha)*gamma[i])/W[i])^((1-alpha)/alpha))

deltaY[i] = (spacing)^(1/(0.5*alpha))*(sin((0.5*alpha)*
gamma[i]))/((cos(gamma[i]))^(1/(0.5*alpha)))*
(cos((1-(0.5*alpha))*gamma[i])/W[i])^((1-(0.5*alpha))/
(0.5*alpha)))

L[i] = sum(deltaX[1:i])
St[i] = sum(deltaY[1:i])
}

#Vasicek with jumps

delta_t = spacing

for(i in 1:(n-1)){
dr = a*(b-r[i])*delta_t + sigma*(L[i+1]-L[i])
r[i+1] = r[i] + dr
}

t = t[2:(n+1)]

#Stochastic integral part

sto_int = c()

```

```

sto_int_s = c()

for(i in 1:(length(t)-1)){
  sto_int[i] = exp(-a*t[i])*(L[i+1]-L[i])
  sto_int_s[i] = sum(sto_int[1:i])
}

sto_int_s

#Main expression

r1 = pmax(r-K,0)

values = c()

for(i in 1:length(t)){
  values[i] = 1/St[i]*r1[i]*sto_int_s[i]*1/sigma
}

print(counter)
counter = counter + 1

process[j,] = values
}

expectation = c()
for(i in 1:n){
  expectation[i] = 1/m*sum(process[,i])
}

setwd("C:/Users/Tor_Martin/Documents")

png("zhang0_8.png", width=16, height=10, units='cm',
    res=1000)

```



```

plot(t[1:(n-1)], expectation[1:(n-1)], type = "l", xlim =
      c(0,t[n]), ylim = c(-5,5), xlab = "Time",
ylab = "Expectation", main = "r0 = 0.048, strike =
      0.05, longterm interest rate = 0.04")

lines(t[1:(n-1)], expectationBM[1:(n-1)], col = "red")

dev.off()

```

Listing A.4: Simulation of figure 3.5 (Brownian motion added to figure 3.4)

```

# Vasicek (Zhang) in the case of Brownian motion

setwd("C:/Users/Tor_Martin/Documents")

# png("zhang1_BM.png", width=16, height=10, units='cm',
      res=1000)

n = 600 # Number of points in Vasicek
K = 0.05 # Strike of the caplet

r0 = 0.048 # Initial interest rate
a = 0.05 # Will depend on monetary policy
b = 0.04 # Theta: Longterm interest rate
sigma = 0.004 # Volatility

m = 10000 # Number of simulations for expectation
counter = 1
process = matrix(0,m,n)
for(j in 1:m){
  r = c()
  r[1] = r0

# Simulation of trajectory for Brownian motion

```

```

sigma_BM = 0.5 #Volatility of Brownian motion
b = 0.06 #Drift of Brownian motion
t = c()
t[1] = 0 #t0 = t[1]
t[n] = 600
spacing = (t[n]-t[1])/(n)

for (i in 2:(n+1)){
t[i] = t[i-1] + spacing
}

N = rnorm(n)
deltaX = c()
L = c()
for(i in 1:n){
deltaX[i] = sigma_BM*N[i]*sqrt(spacing) + b*spacing
L[i] = sum(deltaX[1:i])
}

#Vasicek with BM

delta_t = spacing

for(i in 1:(n-1)){
dr = a*(b-r[i])*delta_t + sigma*(L[i+1]-L[i])
r[i+1] = r[i] + dr
}

t = t[2:(n+1)]

#plot(t,r,type = "l")

#Stochastic integral part

sto_int = c()
sto_int_s = c()

```

```

for(i in 1:(length(t)-1)){
sto_int[i] = exp(-a*t[i])*(L[i+1]-L[i])
sto_int_s[i] = sum(sto_int[1:i])
}

sto_int_s

#Main expression

r1 = pmax(r-K,0)

values = c()

for(i in 1:length(t)){
values[i] = 1/t[i]*r1[i]*sto_int_s[i]*1/sigma
}

print(counter)
counter = counter + 1

process[j,] = values
}

expectationBM = c()
for(i in 1:n){
expectationBM[i] = 1/m*sum(process[,i])
}

setwd("C:/Users/Tor_Martin/Documents")

png("zhang1_BM_final.png", width=16, height=10, units='
cm', res=1000)

```

```

plot(t[1:(n-1)], expectationBM[1:(n-1)], type = "l", xlim
     = c(0,t[n]), ylim = c(-5,5), xlab = "Time", ylab =
     "Expectation")

lines(t[1:(n-1)], expectationBM[1:(n-1)], col = "red")

dev.off()

```

Listing A.5: Simulation of figure 3.6

```

#Endowment policy with jumps

m = 200000      #Number of simulations for expectation
K = 0.05        #Strike of the caplet (should be higher
                then longterm interest rate)
delta = 0.03    #Fixed interest rate over the periode
                of the contract
s = 35          #Age of the individual at the beginning
                of the contract
T = 65          #Age of the individual at the end of
                the contract
k = 12          #Number of observeed interest rate
                changes per year e.g. 12 if change every month
n = (T-s)*k     #Number of points over the total period


r0 = 0.048      #Initial intrest rate
a = 0.05        #Will depend on monetary policy
b = 0.04        #Teta: Longterm intrest rate
sigma = 0.004   #Volatility


#Transition probabilities

aa = -7.85785
bb = 0.01538
cc = 5.77355*10^(-4)

```

```

uxt = function(x){
exp(aa+bb*x+cc*x^2)
}

pxx = function(s,T){
exp(-(integrate(uxt,s,T)$value))
}

counter = 1
process = matrix(0,m,n)
for(j in 1:m){
r = c()
r[1] = r0

#Simulation of trajectory for symmetric alpha stable
process
alpha = 1.9

t = c()
t[1] = 0 #t0 = t[1]
t[n] = (T-s)*365
spacing = (t[n]-t[1])/(n)

for (i in 2:(n+1)){
t[i] = t[i-1] + spacing
}

gamma = runif(n, -pi/2, pi/2)
W = rexp(n)
deltaX =c()
deltaY = c()
L = c()
St = c()

for (i in 1:n){

```

```

deltaX[i] = (spacing)^(1/alpha)*(sin(alpha*gamma[i]))/
    ((cos(gamma[i]))^(1/alpha))*
(cos((1-alpha)*gamma[i])/W[i])^((1-alpha)/alpha)

deltaY[i] = (spacing)^(1/(0.5*alpha))*(sin((0.5*alpha)*
    gamma[i]))/((cos(gamma[i]))^(1/(0.5*alpha)))*
(cos((1-(0.5*alpha))*gamma[i])/W[i])^((1-(0.5*alpha))/
    (0.5*alpha))

L[i] = sum(deltaX[1:i])
St[i] = sum(deltaY[1:i])
}

#Vasicek with jumps

delta_t = spacing

for(i in 1:(n-1)){
dr = a*(b-r[i])*delta_t + sigma*(L[i+1]-L[i])
r[i+1] = r[i] + dr
}

t = t[2:(n+1)]

#Stochastic integral part

sto_int = c()
sto_int_s = c()

for(i in 1:(length(t)-1)){
sto_int[i] = exp(-a*t[i])*(L[i+1]-L[i])
sto_int_s[i] = sum(sto_int[1:i])
}

#sto_int_s

```

```

#Main expression

r1 = pmax(r-K,0)

values = c()

for(i in 1:length(t)){
values[i] = 1/St[i]*r1[i]*sto_int_s[i]*1/sigma
}

print(counter)
counter = counter + 1

process[j,] = values
}

expectation = c()
for(i in 1:n){
expectation[i] = 1/m*sum(process[,i])
}

V0 = exp(-delta*(T-s))*expectation*pxx(s,T)

setwd("C:/Users/Tor_Martin/Documents")

png("endowment_1_8.png", width=16, height=10, units='cm',
    , res=1000)

plot((t[1:(n-1)]/365),V0[1:(n-1)], type = "l", ylim = c
    (-0.2,0.3),xlab = "Time",
ylab = "", main = "Sensitivity_of_an_endowment_policy")

```

```
dev.off()
```

Listing A.6: Simulation of figure 3.7

```
#Endowment policy with Brownian motion

m = 10000          #Number of simulations for expectation
K = 0.05           #Strike of the caplet (should be higher
                   then longterm interest rate)
delta = 0.03       #Fixed interest rate over the periode
                   of the contract
s = 35             #Age of the individual at the beginning
                   of the contract
T = 65             #Age of the individual at the end of
                   the contract
k = 12             #Number of observeed interest rate
                   changes per year e.g. 12 if change every month
n = (T-s)*k        #Number of points over the total period


r0 = 0.048         #Initial intrest rate
a = 0.05           #Will depend on monetary policy
b = 0.04           #Teta: Longterm intrest rate
sigma = 0.004      #Volatility


#Transition probabilities

aa = -7.85785
bb = 0.01538
cc = 5.77355*10^(-4)

uxt = function(x){
  exp(aa+bb*x+cc*x^2)
}

pxx = function(s,T){
  exp(-(integrate(uxt,s,T)$value))
```



```

}

counter = 1
process = matrix(0,m,n)
for(j in 1:m){
r = c()
r[1] = r0

#Simulation of trajectory for Brownian motion

sigma_BM = 0.5 #Volatility of Brownian motion
bbb = 0.06 #Drift of Brownian motion

t = c()
t[1] = 0 #t0 = t[1]
t[n] = (T-s)*365
spacing = (t[n]-t[1])/(n)

for(i in 2:(n+1)){
t[i] = t[i-1] + spacing
}

N = rnorm(n)
deltaX = c()
L = c()
for(i in 1:n){
deltaX[i] = sigma_BM*N[i]*sqrt(spacing) + bbb*spacing
L[i] = sum(deltaX[1:i])
}

#Vasicek with jumps

delta_t = spacing

for(i in 1:(n-1)){

```

```

dr = a*(b-r[i])*delta_t + sigma*(L[i+1]-L[i])
r[i+1] = r[i] + dr
}

t = t[2:(n+1)]

#Stochastic integral part

sto_int = c()
sto_int_s = c()

for(i in 1:(length(t)-1)){
  sto_int[i] = exp(-a*t[i])*(L[i+1]-L[i])
  sto_int_s[i] = sum(sto_int[1:i])
}

#sto_int_s

#Main expression

r1 = pmax(r-K,0)

values = c()

for(i in 1:length(t)){
  values[i] = 1/t[i]*r1[i]*sto_int_s[i]*1/sigma
}

print(counter)
counter = counter + 1

process[j,] = values
}

expectation = c()
for(i in 1:n){

```

```

expectation[i] = 1/m*sum(process[,i])
}

V0 = exp(-delta*(T-s))*expectation*pxx(s,T)

setwd("C:/Users/Tor_Martin/Documents")

png("endowment_BM.png", width=16, height=10, units='cm',
    , res=1000)

plot((t[1:(n-1)]/365),V0[1:(n-1)], type = "l", ylim = c
    (-0.0003,0.0042),xlab = "Time",
ylab = "", main = "Sensitivity_of_an_endowment_policy_
    with_BM")

#lines(t[1:(n-1)],expectationBM[1:(n-1)],col = "red")

dev.off()

```

Listing A.7: Simulation of figure 3.8

```

#Term life insurance with jumps

m = 100000          #Number of simulations for
                    expectation
delta = 0.03        #Fixed interest rate over the periode
                    of the contract
s = 35              #Age of the individual at the beginning
                    of the contract
T = 65              #Age of the individual at the end of
                    the contract
k = 12              #Number of obsereved interest rate
                    changes per year e.g. 12 if change every month

```

```

n = (T-s)*k      #Number of points over the total period

x0 = 6           #Initial value of stock
sigma = 0.5      #Volatility
G = 5            #Guarantee
K = 8            #Strike price

#Transition probabilities

aa = -7.85785
bb = 0.01538
cc = 5.77355*10^(-4)

uxt = function(x){
exp(aa+bb*x+cc*x^2)
}

pxx = function(s,T){
exp(-(integrate(uxt,s,T)$value))
}

counter = 1
process = matrix(0,m,n)
for(j in 1:m){

#Simulation of trajectory for symmetric alpha stable
process
alpha = 1.9

t = c()
t[1] = 0          #t0 = t[1]
t[n] = (T-s)*365
spacing = (t[n]-t[1])/(n)

for (i in 2:(n+1)){

```

```

t[i] = t[i-1] + spacing
}

gamma = runif(n, -pi/2, pi/2)
W = rexp(n)
deltaX = c()
deltaY = c()
L = c()
St = c()

for (i in 1:n){
deltaX[i] = (spacing)^(1/alpha)*(sin(alpha*gamma[i]))/
((cos(gamma[i]))^(1/alpha))*
(cos((1-alpha)*gamma[i])/W[i])^((1-alpha)/alpha))

deltaY[i] = (spacing)^(1/(0.5*alpha))*(sin((0.5*alpha)*
gamma[i]))/((cos(gamma[i]))^(1/(0.5*alpha)))*
(cos((1-(0.5*alpha))*gamma[i])/W[i])^((1-(0.5*alpha))/
(0.5*alpha)))

L[i] = sum(deltaX[1:i])
St[i] = sum(deltaY[1:i])
}

L = L/max(abs(L))
St = St/max(abs(L))

t = t[2:(n+1)]

#Stochastic integral part

sto_int = c()
sto_int_s = c()

for(i in 1:(length(t)-1)){
sto_int[i] = exp(sigma*L[i])*(L[i+1]-L[i])

```

```

sto_int_s[i] = sum(sto_int[1:i])
}

#sto_int_s

#Main expression

r1 = pmax(K-x0*exp(sigma*L),G)

values = c()

for(i in 1:length(t)){
values[i] = 1/St[i]*r1[i]*sto_int_s[i]*1/sigma
}

print(counter)
counter = counter + 1

process[j,] = values
}

expectation = c()
for(i in 1:n){
expectation[i] = 1/m*sum(process[,i])
}

delta_t = spacing

int_V0 = c()
V0 = c()

for(i in 1:length(t)){
int_V0[i] = exp(-delta*(t[i]/365))*expectation[i]*pxx(s
,s+(t[i]/365))*uxt(s+t[i]/365)*(delta_t/365)
V0[i] = sum(int_V0[1:i])
}

```

```

setwd("C:/Users/Tor_Martin/Documents")

png("term_life6_1_9.png", width=16, height=10, units='
  cm', res=1000)

plot((t[1:(n-1)]/365), V0[1:(n-1)], type = "l", xlab = "
  Time",
ylab = "", main = "")

dev.off()

```

Listing A.8: Simulation of figure 3.9

```

#Term life insurance with Brownian motion

m = 100000      #Number of simulations for expectation
delta = 0.03    #Fixed interest rate over the periode
               of the contract
s = 35          #Age of the individual at the beginning
               of the contract
T = 65          #Age of the individual at the end of
               the contract
k = 12          #Number of obsereved interest rate
               changes per year e.g. 12 if change every month
n = (T-s)*k     #Number of points over the total period

x0 = 20         #Initial value of stock
sigma = 0.5     #Volatility
G = 20          #Guarantee
K = 50          #Strike price

#Transition probabilities

aa = -7.85785

```

```

bb = 0.01538
cc = 5.77355*10^(-4)

uxt = function(x){
exp(aa+bb*x+cc*x^2)
}

pxx = function(s,T){
exp(-(integrate(uxt,s,T)$value))
}

counter = 1
process = matrix(0,m,n)
for(j in 1:m){

#Simulation of trajectory for Brownian motion

sigma_BM = 0.5 #Volatility of Brownian motion
bbb = 0.1 #Drift of Brownian motion

t = c()
t[1] = 0 #t0 = t[1]
t[n] = (T-s)*365
spacing = (t[n]-t[1])/(n)

for (i in 2:(n+1)){
t[i] = t[i-1] + spacing
}

N = rnorm(n)
deltaX = c()
L = c()
for(i in 1:n){
deltaX[i] = sigma_BM*N[i]*sqrt(spacing) + bbb*spacing
L[i] = sum(deltaX[1:i])
}

```



```

L = L/(max(abs(L)))

t = t[2:(n+1)]

#Stochastic integral part

sto_int = c()
sto_int_s = c()

for(i in 1:(length(t)-1)){
  sto_int[i] = exp(sigma*L[i])*(L[i+1]-L[i])
  sto_int_s[i] = sum(sto_int[1:i])
}

#sto_int_s

#Main expression

r1 = pmax(K-x0*exp(sigma*L),G)

values = c()

for(i in 1:length(t)){
  values[i] = 1/t[i]*r1[i]*sto_int_s[i]*1/sigma
}

print(counter)
counter = counter + 1

process[j,] = values
}

expectation = c()
for(i in 1:n){
  expectation[i] = 1/m*sum(process[,i])
}

```

```

}

delta_t = spacing

int_V0 = c()
V0 = c()

for(i in 1:length(t)){
  int_V0[i] = exp(-delta*(t[i]/365))*expectation[i]*pxx(s
    ,s+(t[i]/365))*uxt(s+t[i]/365)*(delta_t/365)
  V0[i] = sum(int_V0[1:i])
}

#setwd("C:/Users/Tor Martin/Documents")

#png("term_life_BM.png", width=16, height=10, units='cm'
  , res=1000)

plot((t[1:(n-1)]/365),V0[1:(n-1)], type="l",xlab="
  Time",
ylab="", main="")

#lines(t[1:(n-1)],expectationBM[1:(n-1)], col="red")

#dev.off()

```

A.3 Bibliography

Bibliography

- [1] Benth, Fred E.: Option Theory with Stochastic Analysis: An Introduction to Mathematical Finance. Springer (2004).
- [2] Brigo, D., Mercurio, F.: Interest Rate Models - Theory and Practice 2nd edition. Springer (2006).
- [3] Cont, R., Tankov, R.: Financial modelling with jump processes. Chapman & Hall (2004).
- [4] Di Nunno, Øksendal, B., Proske, F.: Malliavin Calculus for Lévy Processes with application to Finance. Springer (2008).
- [5] Evans, Lawrence C.: Partial Differential Equations, 2nd edition. American Mathematical Society (2010).
- [6] Fournié, E. Lasry, J.-M., Lebuchoux, J., Lions, P.-L, Touzi, N.: Applications of Malliavin Calculus to Monte Carlo methods in finance. *Finance Stoch.*, 3(4) : 391 – 412(1999).
- [7] Fournié, E. Lasry, J.-M., Lebuchoux, J., Lions, P.-L, Touzi, N.: Applications of Malliavin Calculus to Monte Carlo methods in finance II . *Finance Stoch.*, 5(2) : 201 – 236(2001).
- [8] Haadem, S. ,Proske, F.: On the construction and Malliavin differentiability of solutions of Lévy noise driven SDE's with singular coefficients. *Journal of functional analysis*, 266, 5321 – 5359 (2014).
- [9] Ikeda, N., Watanabe, S.: Stochastic Differential Equations and Diffusion Processes, 2nd edition. North-Holland Mathematical Library (1989).

- [10] Jäckel, P.: Monte Carlo Methods in Finance. John Wiley & Sons, Chichester (2002).
- [11] Koller, M.: Stochastic Models in Life Insurance. Springer (2011).
- [12] Meneoukeu-Pamen, O., Meyer-Brandis, T., Nilssen, T., Proske, F., Zhang, T.: A variational approach to the construction and Malliavin differentiability of strong solutions of SDE's. Springer (2013).
- [13] Peszat, S., Zabczyk, J.: Stochastic Partial Differential Equations with Lévy Noise. Cambridge university press (2007).
- [14] Protter, P.: Stochastic integration and differential equations, 2nd edition, Springer (2005).
- [15] Rynne, B., Youngson, M.: Linear Functional Analysis, 2nd edition. Springer (2008).
- [16] Veretennikov, A. Yu.: On strong solutions of Itô equations with jumps. Theory Probab. Appl., 32(1), 148 – 152(1985).
- [17] Zhang, X.: Derivative formulas and gradient estimates for SDE's driven by α -stable processes. Stochastic Processes and Their Applications 123 p.1213-1228 (2012).
- [18] Øksendal, B.: Stochastic Differential Equations. Springer (2010).
- [19] Øksendal, B., Sulem, A.: Applied Stochastic Control of Jump Diffusions. Springer (2007).